

US010691133B1

(12) **United States Patent**
Abeloe

(10) **Patent No.:** **US 10,691,133 B1**
(45) **Date of Patent:** **Jun. 23, 2020**

(54) **ADAPTIVE AND INTERCHANGEABLE NEURAL NETWORKS**

FOREIGN PATENT DOCUMENTS

(71) Applicant: **Apex Artificial Intelligence Industries, Inc.**, Centreville, VA (US)

CN	101763033	12/2009
CN	102522945	6/2014
CN	107272705	10/2017

(72) Inventor: **Kenneth Austin Abeloe**, Carlsbad, CA (US)

OTHER PUBLICATIONS

(73) Assignee: **Apex Artificial Intelligence Industries, Inc.**, Centreville, VA (US)

“The Internet of Things: Opportunities and Applications across Industries—Discussion Summary” International Institute for Analytics, 8 pages (Dec. 2015).

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 0 days.

(Continued)

(21) Appl. No.: **16/734,074**

Primary Examiner — Calvin Cheung

(22) Filed: **Jan. 3, 2020**

(74) *Attorney, Agent, or Firm* — Knobbe, Martens, Olson & Bear LLP

Related U.S. Application Data

(60) Provisional application No. 62/940,762, filed on Nov. 26, 2019.

(51) **Int. Cl.**
G05D 1/02 (2020.01)
G06N 3/04 (2006.01)
G05D 1/00 (2006.01)

(57) **ABSTRACT**

(52) **U.S. Cl.**
CPC **G05D 1/0221** (2013.01); **G05D 1/0088** (2013.01); **G05D 1/0278** (2013.01); **G06N 3/0445** (2013.01)

Methods and systems that allow neural network systems to maintain or increase operational accuracy while being able to operate in various settings. A set of training data is collected over each of at least two different settings. Each setting has a set of characteristics. Examples of setting characteristic types can be time, geographical location, and/or weather condition. Each set of training data is used to train a neural network resulting in a set of coefficients. For each setting, the setting characteristics are associated with the corresponding neural network having the resulting coefficients and neural network structure. A neural network, having the coefficients and neural network structure resulted after training using the training data collected over a setting, would yield optimal results when operated in/under the setting. A database management system can store information relating to, for example, the setting characteristics, neural network coefficients, and/or neural network structures.

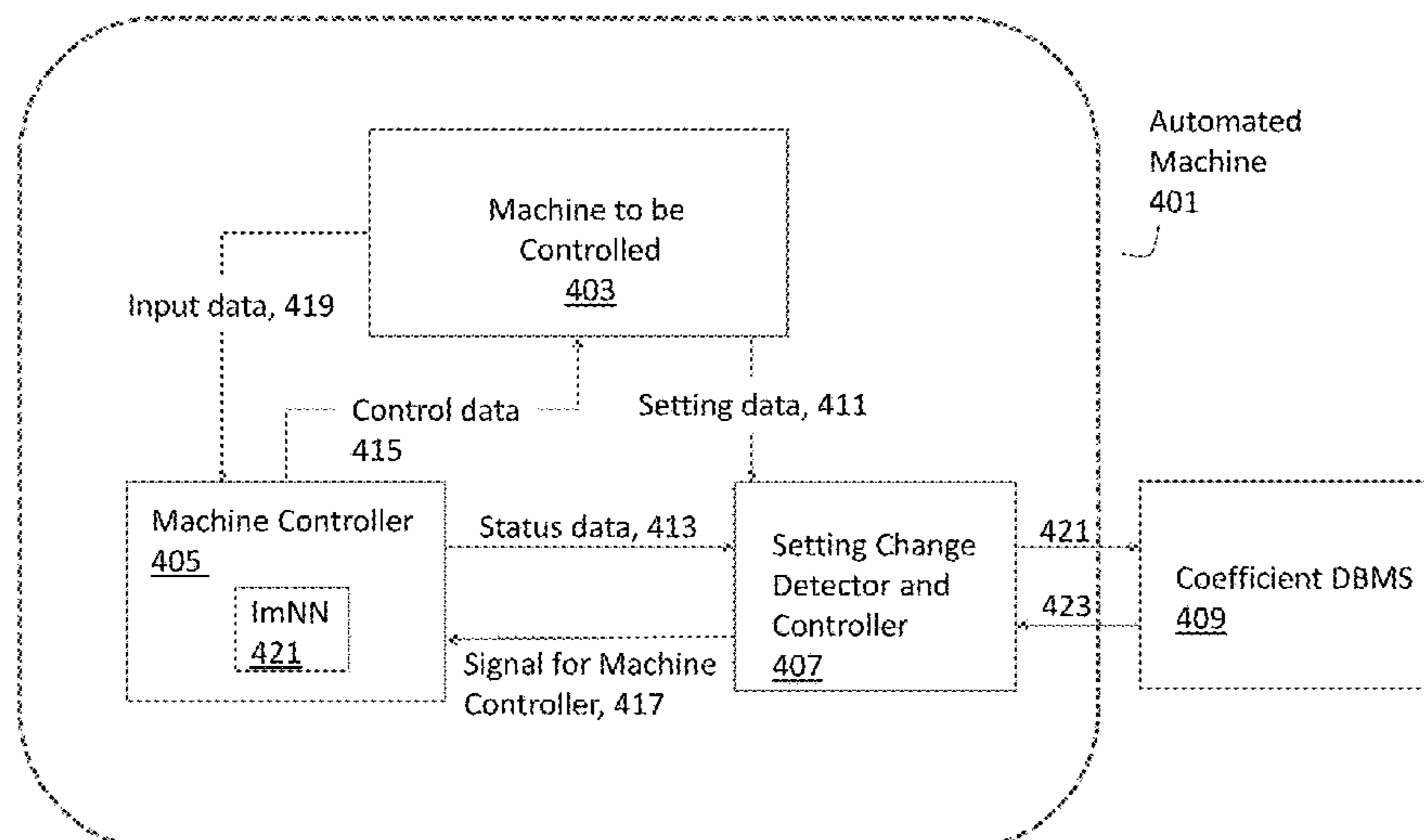
(58) **Field of Classification Search**
CPC .. G05D 1/0221; G05D 1/0088; G05D 1/0278; G06N 3/0445
See application file for complete search history.

(56) **References Cited**

U.S. PATENT DOCUMENTS

5,048,100 A 9/1991 Kuperstein
5,179,631 A 1/1993 Guddanti et al.
(Continued)

30 Claims, 13 Drawing Sheets



(56)

References Cited

U.S. PATENT DOCUMENTS

5,255,347	A *	10/1993	Matsuba	G06N 3/0454 706/25
5,668,926	A	9/1997	Karaali		
5,845,271	A	12/1998	Thaler		
5,870,729	A *	2/1999	Yoda	G06K 9/6272 706/26
6,574,754	B1	6/2003	Smith		
6,601,051	B1 *	7/2003	Lo	H03H 21/0012 706/23
7,010,513	B2	3/2006	Tamura		
7,063,476	B1	6/2006	Mims		
7,800,490	B2	9/2010	Allen et al.		
7,979,173	B2	7/2011	Breed		
8,677,071	B2	3/2014	Gibson et al.		
8,761,991	B1	6/2014	Ferguson et al.		
8,775,341	B1	7/2014	Commons		
8,782,653	B2	7/2014	Gibson et al.		
8,892,485	B2	11/2014	Aparin et al.		
9,025,140	B2	5/2015	Zhu et al.		
9,460,382	B2 *	10/2016	Canoy	G06N 3/049
9,875,440	B1	1/2018	Commons		
10,347,127	B2	9/2019	Droz et al.		
10,444,754	B2	10/2019	Fairfield et al.		
10,451,712	B1 *	10/2019	Madhow	G01S 7/417
2003/0217021	A1 *	11/2003	Jacobson	F02D 41/1405 706/16
2009/0276385	A1 *	11/2009	Hill	G06N 3/08 706/25
2010/0194593	A1	8/2010	Mays		
2012/0259804	A1 *	10/2012	Brezzo	G06N 3/049 706/25
2013/0073497	A1 *	3/2013	Akopyan	G06N 3/063 706/27
2013/0141233	A1	6/2013	Jacobs et al.		
2013/0231824	A1	9/2013	Wilson et al.		
2013/0325776	A1 *	12/2013	Ponulak	G06N 3/08 706/25
2014/0156568	A1 *	6/2014	Ganguly	G06Q 10/0637 706/12
2015/0042491	A1	2/2015	Burnison		
2017/0308800	A1	10/2017	Chichon et al.		
2018/0025268	A1	1/2018	Teig et al.		
2018/0246853	A1 *	8/2018	Fowers	G06F 9/3891
2018/0259345	A1 *	9/2018	Wang	G06F 3/00
2018/0268291	A1 *	9/2018	Turner	G06N 3/0427
2018/0300556	A1 *	10/2018	Varekar	G06K 9/00711
2018/0307950	A1 *	10/2018	Nealis	G06K 9/66
2018/0322365	A1 *	11/2018	Yehezkel Rohekar	G06K 9/00986
2018/0322382	A1 *	11/2018	Mellempudi	G06N 3/0454
2019/0087726	A1 *	3/2019	Greenblatt	G06N 3/08
2019/0138829	A1	5/2019	Gurghian et al.		
2019/0147298	A1 *	5/2019	Rabinovich	G06N 3/0472 382/157
2019/0156180	A1 *	5/2019	Nomura	G06N 3/0445
2019/0176841	A1	6/2019	Englard et al.		
2019/0311597	A1 *	10/2019	Sohn	G08B 21/0469
2019/0347567	A1 *	11/2019	Njie	G16B 20/20
2019/0392268	A1 *	12/2019	Tariq	G06K 9/3233
2020/0033868	A1 *	1/2020	Palanisamy	G05D 1/0221
2020/0034197	A1 *	1/2020	Nagpal	G06F 9/50
2020/0043197	A1 *	2/2020	Bazin	G06N 20/20

OTHER PUBLICATIONS

Bengio et al., "Greedy Layer-Wise Training of Deep Networks," University of Montreal, 8 pages.

Broomhead et al., "Radial Basis Functions, Multi-Variable Functional Interpolation and Adaptive Networks," Royal Signals & Radar Establishment, RSRE Memorandum No. 4148, 39 pages (Mar. 28, 1988).

Chang et al., "CLKN: Cascaded Lucas-Kanade Networks for Image Alignment," 9 pages (2017).

Chen et al., "Big Data Deep Learning: Challenges and Perspectives," *IEEE Access*, vol. 2, pp. 514-525 (2014).

Chen et al., "DeepDriving: Learning Affordance for Direct Perception in Autonomous Driving," Princeton University paper, 9 pages.

Chung et al., "Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling," University of Montreal, arXiv:1412.3555v1, 9 pages (Dec. 11, 2014).

Cortes et al., "Support-Vector Networks," *Machine Learning*, 20, pp. 273-297 (1995).

Dai et al., "Dynamic Learning from Adaptive Neural Network Control of a Class of Nonaffine Nonlinear Systems," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 25, No. 1 (Jan. 2014).

Desai et al., "A Survey of Image Registration Techniques Using Neural Networks," *Int'l Journal of Computer Sciences & Engineering*, vol. 3(12), pp. 57-60 (Dec. 2015).

Dixit et al., "Autonomous Vehicles: Disengagements, Accidents and Reaction Times," *Plos One*, 15 pages (Dec. 20, 2016).

Durrant-Whyte et al., "Simultaneous Localization and Mapping: Part I," *IEEE Robotics & Automation Magazine*, 10 pages (Jun. 2006).

Durrant-Whyte et al., "Simultaneous Localization and Mapping: Part II," *IEEE Robotics & Automation Magazine*, 10 pages (Sep. 2006).

Elman, J.L., "Finding Structure in Time," *Cognitive Science* 14, pp. 179-211 (1990).

Eykholt et al., "Robust Physical-World Attacks on Deep Learning Visual Classification," arXiv:1707.08945v5, 11 pages (Apr. 10, 2018).

Fischer et al. "Descriptor Matching with Convolution Neural Networks: a Comparison to SIFT," arXiv:1405.5769v1, 10 pages (May 22, 2014).

Florian, R.V., "Reinforcement learning through modulation of spike-timing-dependent synaptic plasticity," Center for Cognitive and Neural Studies, Romania, 36 pages.

Freuder E.C., "Explaining Ourselves: Human-Aware Constraint Reasoning," Proceedings of the 31st AAAI Conference on Artificial Intelligence (AAAI-17), pp. 4858-4862 (2017).

Gao et al., "Unsupervised learning to detect loops using deep neural networks for visual SLAM system," *Auton Robot* 41, pp. 1-18 (2017).

Goodfellow et al., "Explaining and Harnessing Adversarial Examples," published as a conference paper at ICLR 2015, ArXiv:1412.6572v3, 11 pages (Mar. 20, 2015).

Goodfellow et al., "Generative Adversarial Nets," University of Montreal, arXiv:1406.2661v1, 9 pages (Jun. 10, 2014).

Graves et al., "Neural Turing Machines," Google DeepMind, London, United Kingdom, arXiv:1410.5401v2, 26 pages (Dec. 10, 2014).

Han et al., "A structure optimisation algorithm for feedforward neural network construction," *Neurocomputing*, 99, pp. 347-357 (2013).

He et al., "Deep Residual Learning for Image Recognition," arXiv:1512.03385v1, 12 pages (Dec. 10, 2015).

Hinton et al., "Learning and Relearning in Boltzmann Machines," Chapter 7, *Basic Mechanics*, pp. 282-317.

Hopfield, J.J., "Neural Networks and Physical Systems with Emergent Collective Computational Abilities," *PNAS* vol. 79, pp. 2554-2558 (Apr. 1982).

Huval et al., "An Empirical Evaluation of Deep Learning on Highway Driving," arXiv:154.01716v3, 7 pages (Apr. 17, 2015).

Jaeger et al., "Harnessing Nonlinearity: Predicting Chaotic Systems and Saving Energy in Wireless Communication," *Science* 304, 78, 4 pages (2004).

Kim, G.K., "Deep Learning," book review, *Healthcare Informatics Research*, 2016 The Korean Society of Medical Informatics, vol. 22, No. 4. pp. 351-354 (2016).

Kingma et al., "Auto-Encoding Variational Bayes," arXiv:1312.6114v10, 14 pages (May 1, 2014).

Kohonen, T., "Essentials of the self-organizing map," *Neural Networks* 37, pp. 52-65 (2013).

Kohonen, T., "Self-Organized Formation of Topologically Correct Feature Maps," *Biol. Cybern.* 43, pp. 59-69 (1982).

(56)

References Cited

OTHER PUBLICATIONS

Kulkarni et al., “Deep Convolutional Inverse Graphics Network,” arXiv:1503.03167v4, 10 pages (Jun. 22, 2015).

LeCun et al., “Deep Learning—Review” *Nature*, vol. 521, pp. 436-444 (May 28, 2015).

LeCun et al., “Gradient-Based Learning Applied to Document Recognition,” *Proc. Of the IEEE*, pp. 1-46 (Nov. 1998).

Libelium Smart World, 1 page, downloaded Jun. 2, 2018, www.libelium.com.

Ma et al., “Multimodal Convolutional Neural Networks for Matching Image and Sentence,” ICCV paper, pp. 2623-2631 (2015).

Maass et al., “Real-Time Computing Without Stable States: A New Framework for Neural Computation Based on Perturbations,” *Neural Computation* 14, pp. 2531-2560 (2002).

Mehri et al., “SampleRNN: An Unconditional End-to-End Neural Audio Generation Model,” published as a conference paper at ICLR 2017, arXiv:1612.07837v2, 11 pages (Feb. 11, 2017).

Nasr et al., “Neural network control of nonlinear dynamic systems using hybrid algorithm—Review article,” *Applied Soft Computing* 24, pp. 423-431 (2014).

Oh et al., “A design of granular-oriented self-organizing hybrid fuzzy polynomial neural networks,” *Neurocomputing*, 119, pp. 292-307 (2013).

Pecevski et al., “Probabilistic Inference in General Graphical Models through Sampling in Stochastic Networks of Spiking Neurons,” *PLoS Computational Biology*, vol. 7, Issue 12, 1-25 (Dec. 2011).

Rahman et al., “Emerging Technologies with Disruptive Effects: A Review,” *Perintis eJournal*, vol. 7, No. 2, pp. 111-128 (2017).

Ranzato et al., “Efficient Learning of Sparse Representations with an Energy-Based Model,” Courant Institute of Mathematical Sciences, New York University, 8 pages.

Rosenblatt, F., “The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain,” *Psychological Review*, vol. 65, No. 6, pp. 386-408 (1958).

Schuster et al., “Bidirectional Recurrent Neural Networks,” *IEEE Transactions on Signal Processing*, vol. 45, No. 11, pp. 2673-2681 (Nov. 1997).

Seff et al., “Learning from Maps: Visual Common Sense for Autonomous Driving,” arXiv:1611.08583v1, 7 pages (Nov. 25, 2016).

Smolensky, P., “Information Processing in Dynamical Systems: Foundations of Harmony Theory,” University of Colorado at Boulder, Department of Computer Science, 56 pages (Feb. 1986).

Svetlov et al., “A New Implementation of the Algorithm of Adaptive Construction of Hierarchical Neural Network Classifiers,” *Optical Memory and Neural Networks (Information Optics)*, vol. 24, No. 4, pp. 288-294 (2015).

Svetlov et al., “Development of the Algorithm of Adaptive Construction of Hierarchical Neural Network Classifiers,” *Optical Memory and Neural Networks (Information Optics)*, vol. 26, No. 1, pp. 40-46 (2017).

Szegedy et al., “Intriguing properties of neural networks,” arXiv:1312.6199v4, 10 pages (Feb. 19, 2014).

Wang, Y., “Deep Learning in Real Time—Inference Acceleration and Continuous Training,” 17 pages, downloaded Dec. 28, 2017, [https://medium.com/@Synced/deep-learning-in-real-time-inference-acceleration-and-cont . . .](https://medium.com/@Synced/deep-learning-in-real-time-inference-acceleration-and-cont...)

Zeiler et al., “Deconvolutional Networks,” Dept. of Computer Science, Courant Institute, New York University, 8 pages.

Zhou et al., “VoxelNet: End-to-End Learning for Point Cloud Based 3D Object Detection,” arXiv:1711.06396v1, 10 pages (Nov. 17, 2017).

Chang et al., “Interchangeable Hebbian and Anti-Hebbian STDP Applied to Supervised Learning in Spiking Neural Network,” 2018 IEEE International Electron Devices Meeting (IEDM), pp. 15.5.1-15.5.4, Dec. 2018.

* cited by examiner

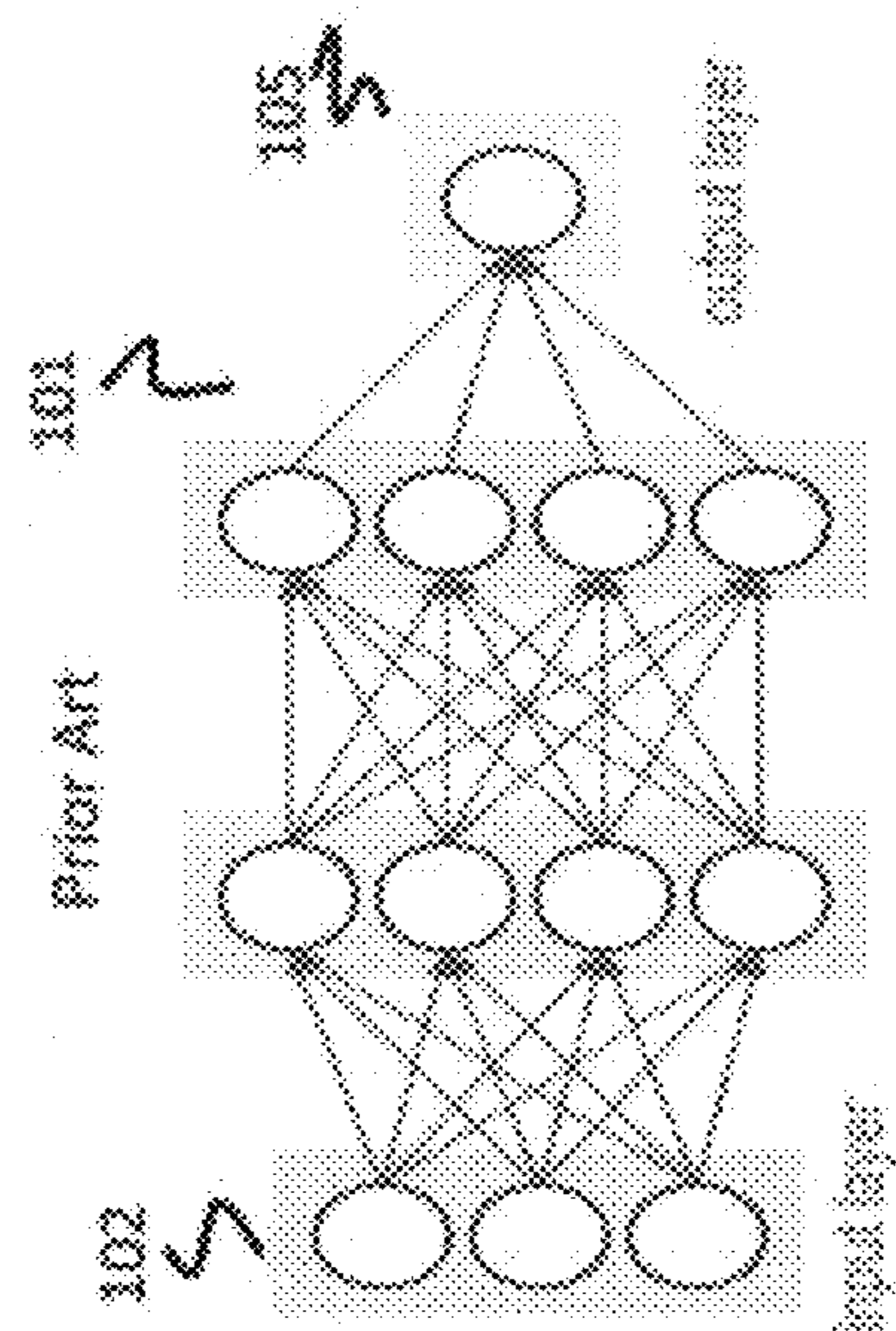


Figure 1

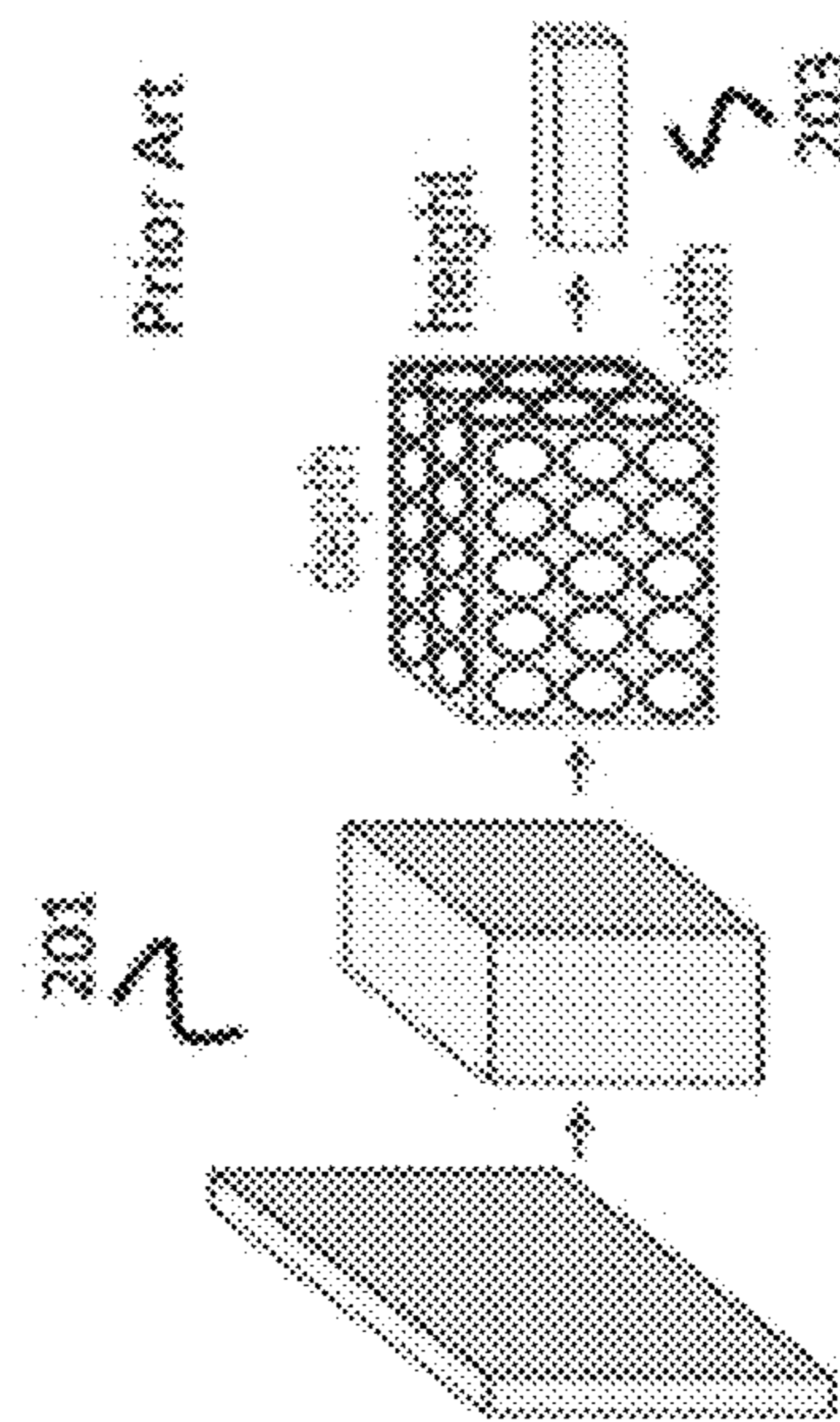


Figure 2

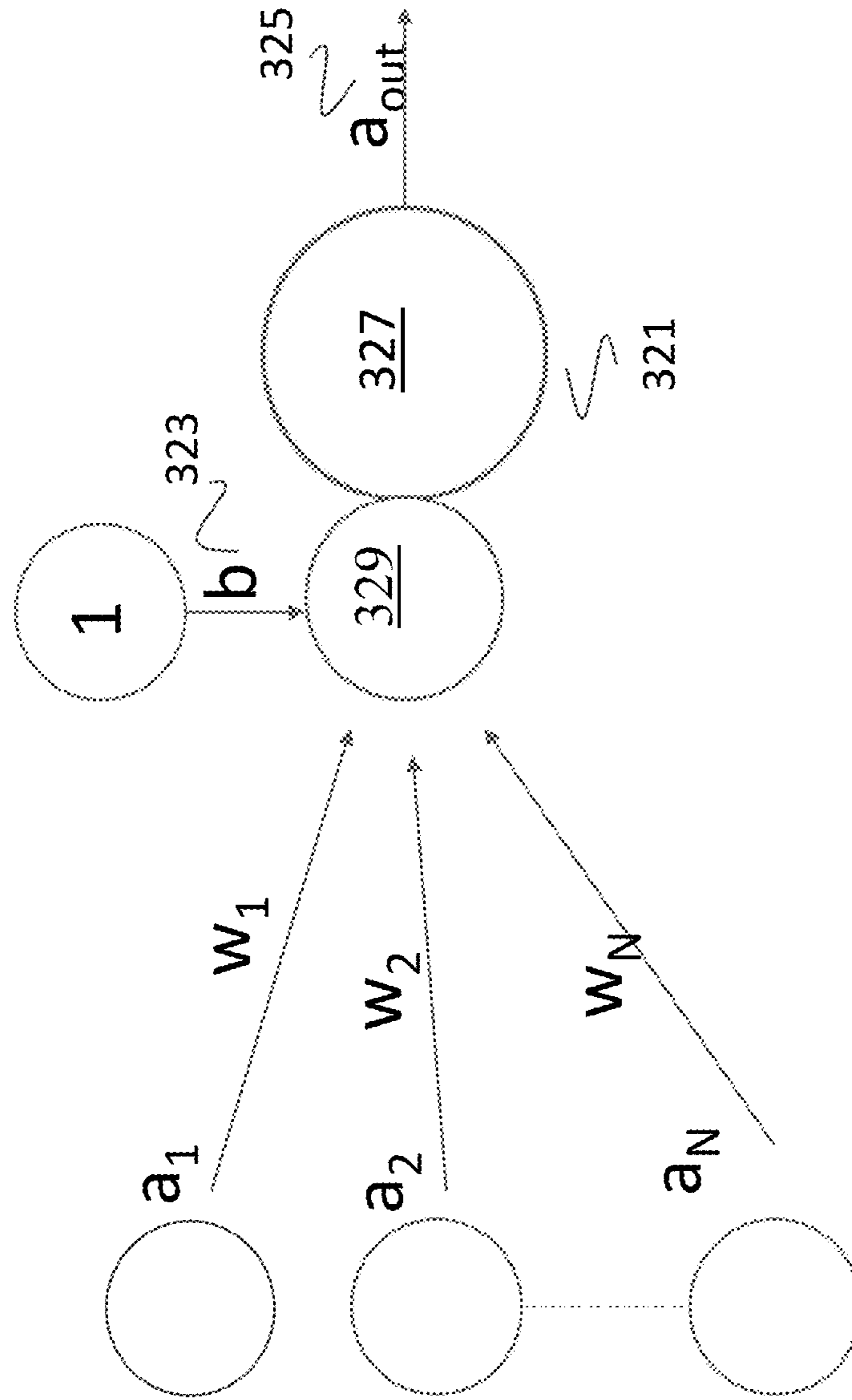


Figure 3

Figure 4

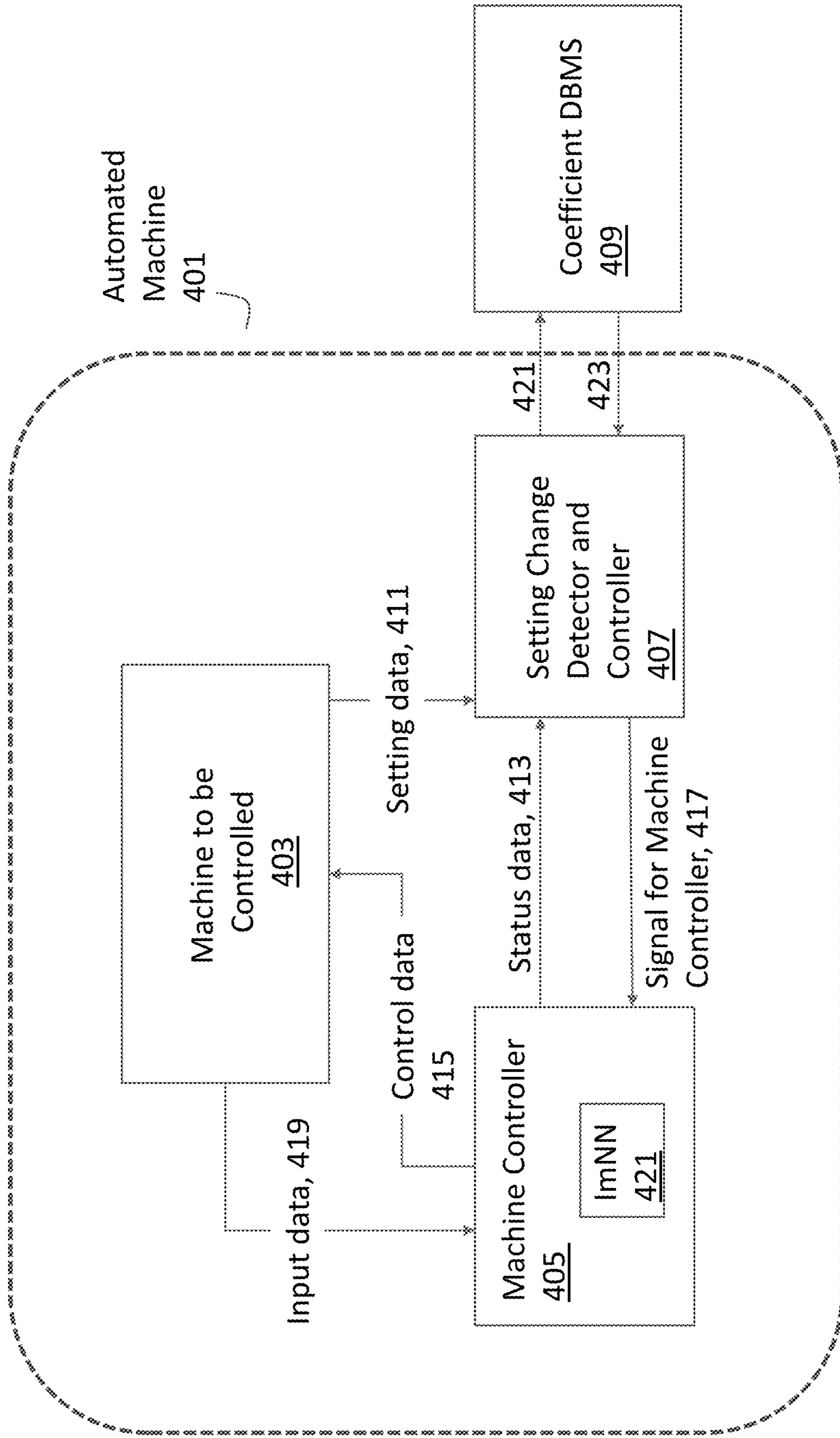


Figure 4a

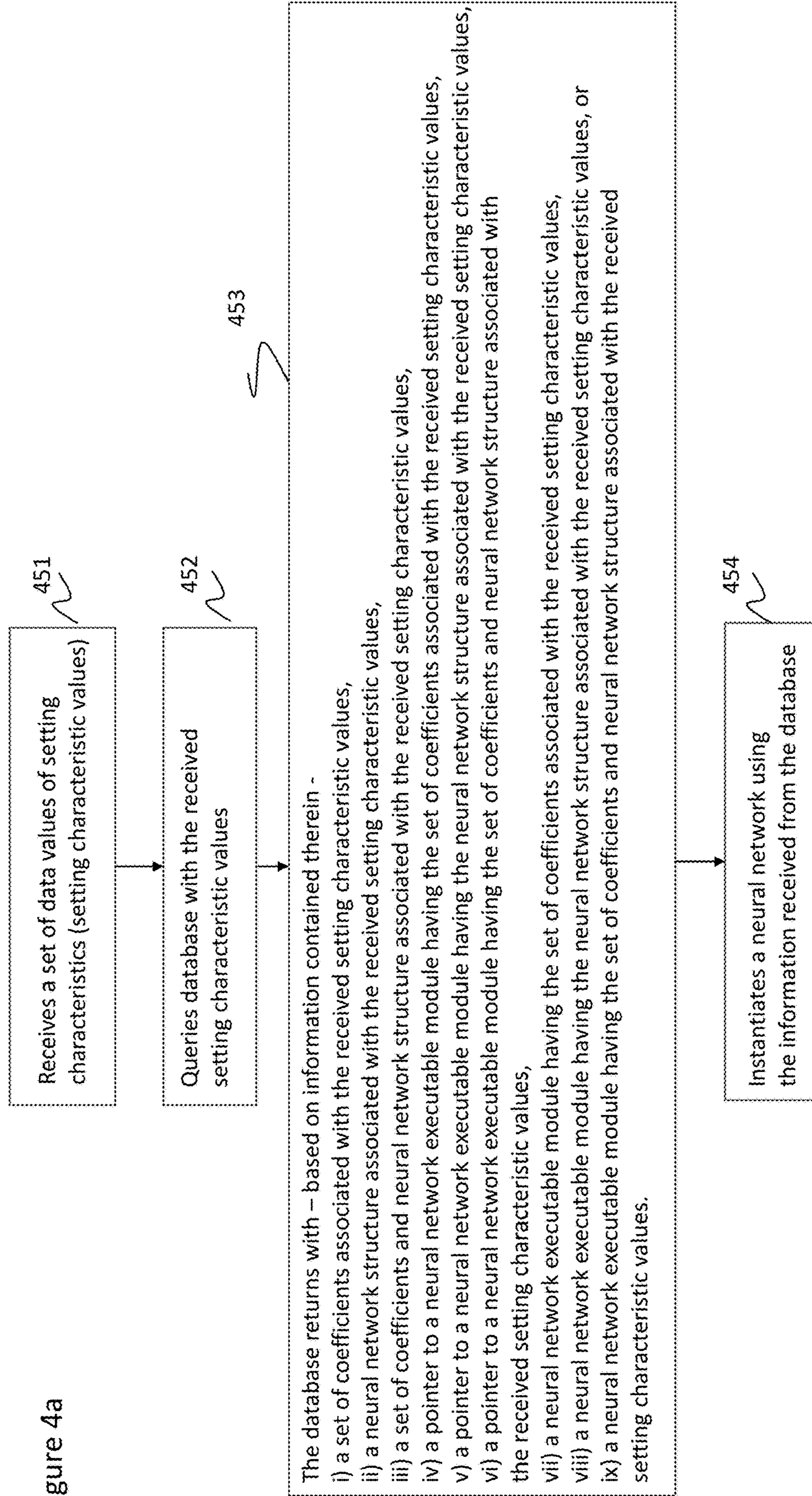


Figure 4b

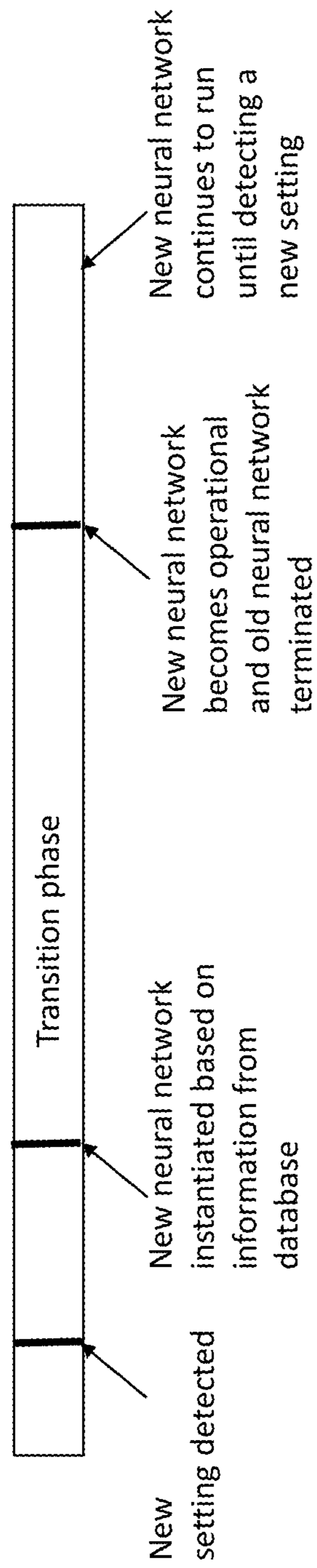
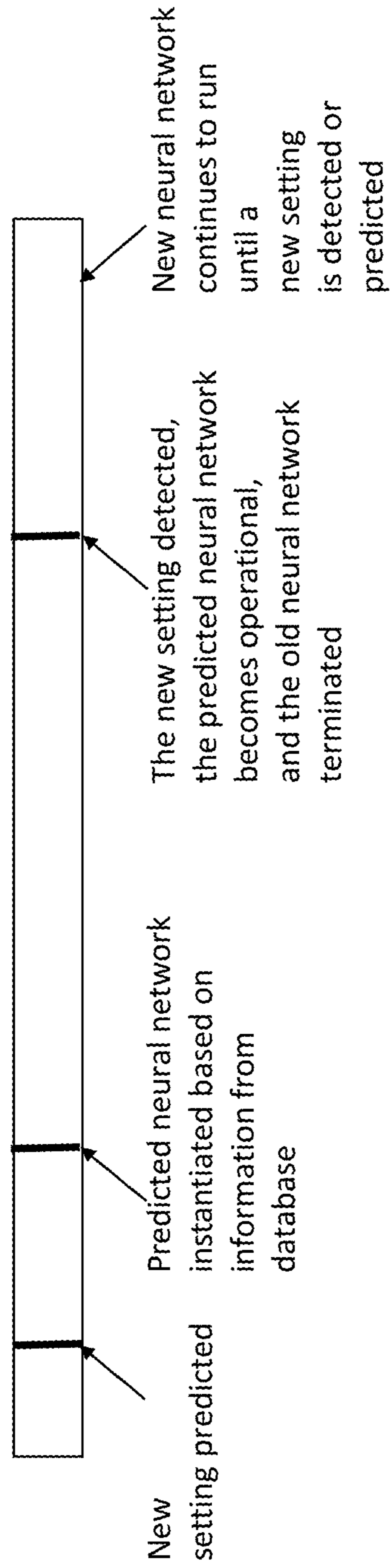


Figure 4c



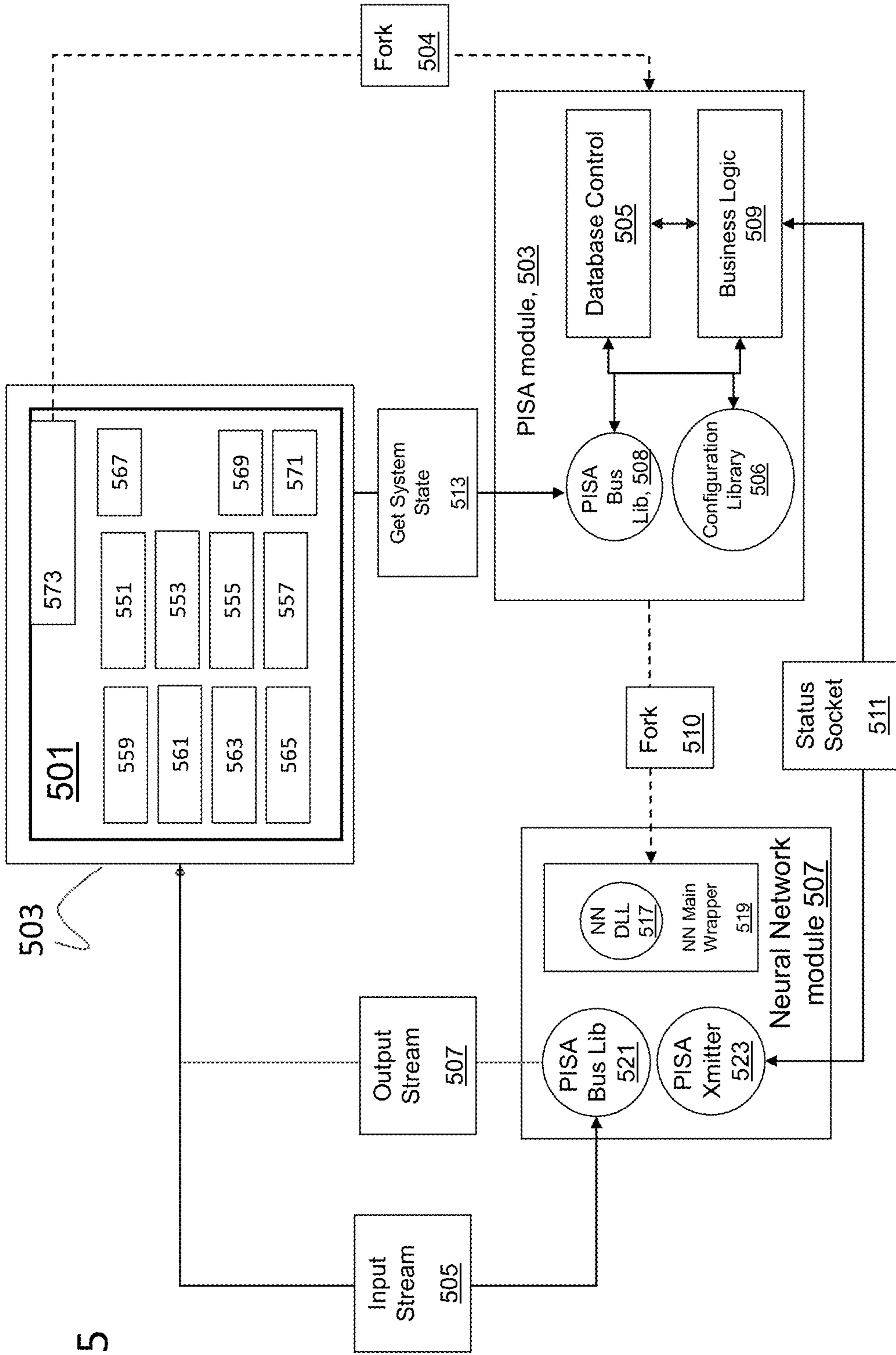


Figure 5

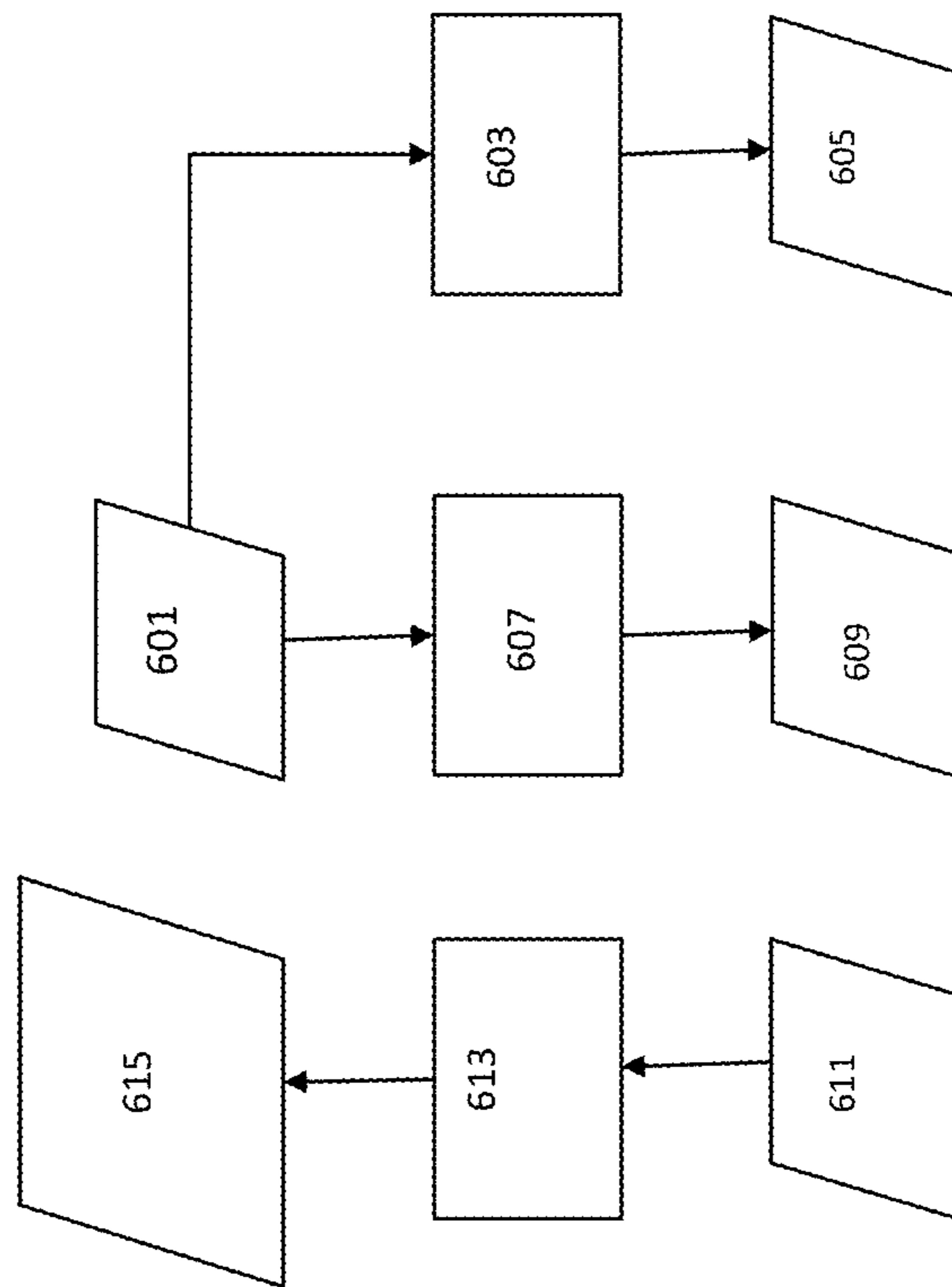


Figure 6

Figure 7

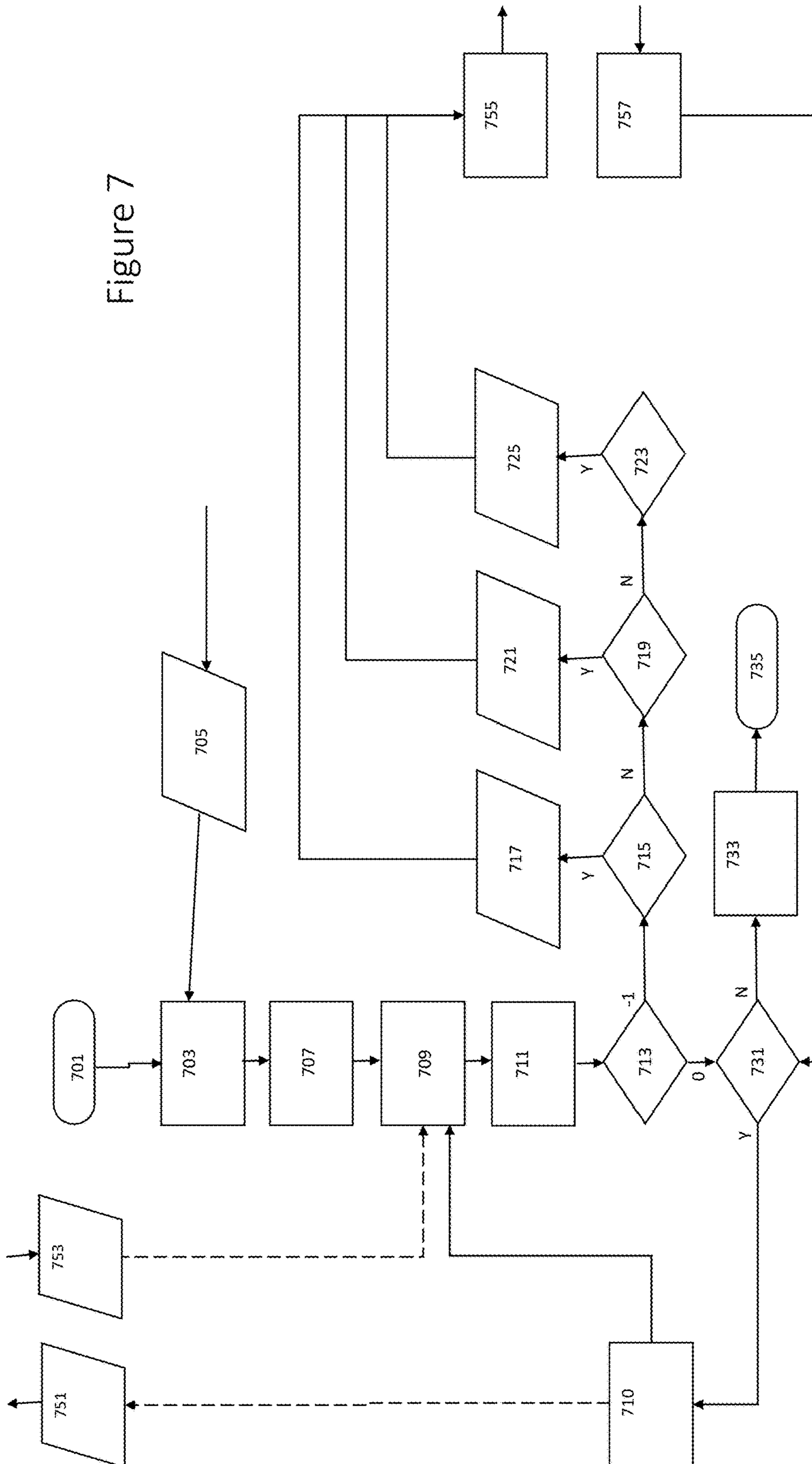
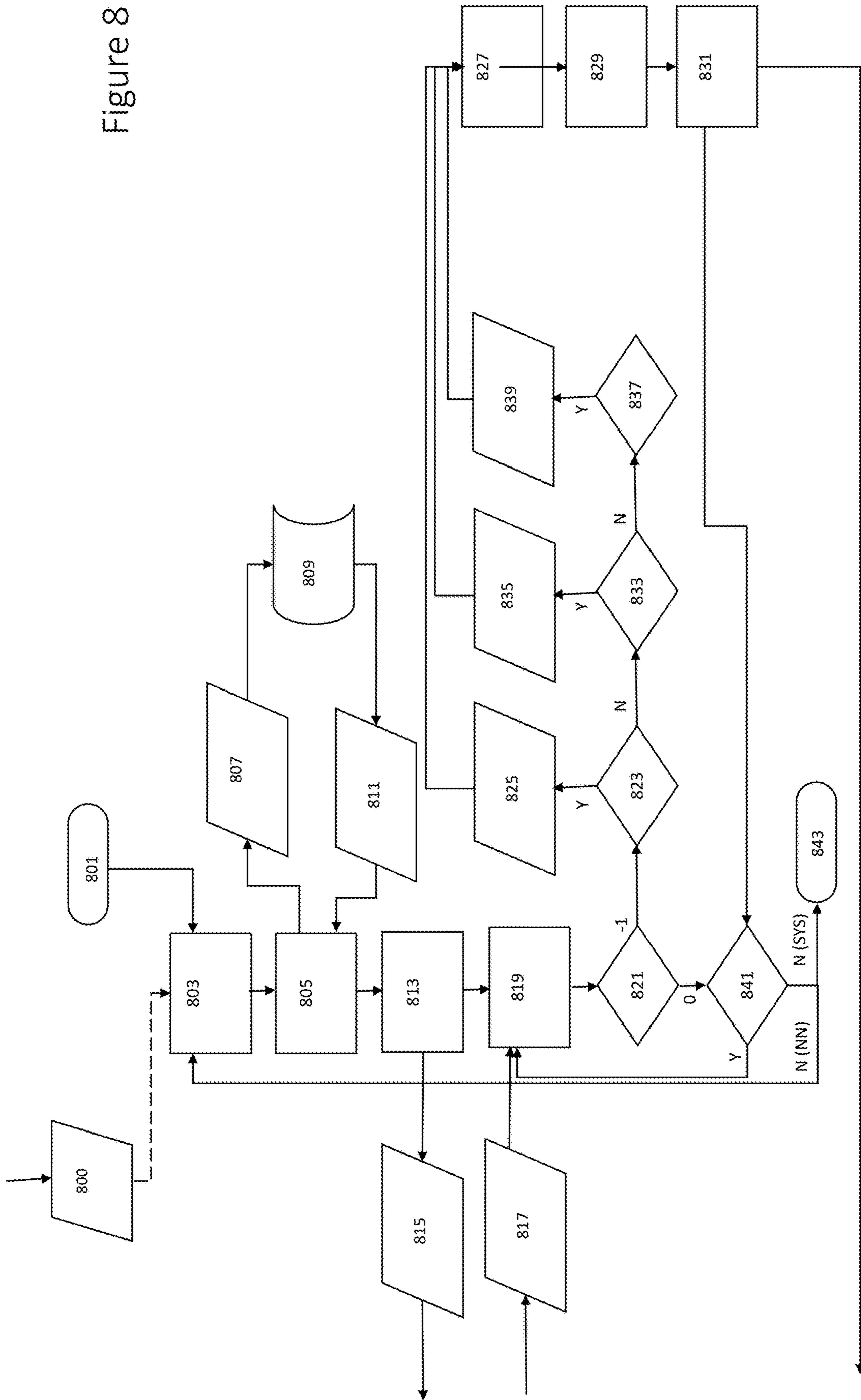


Figure 8



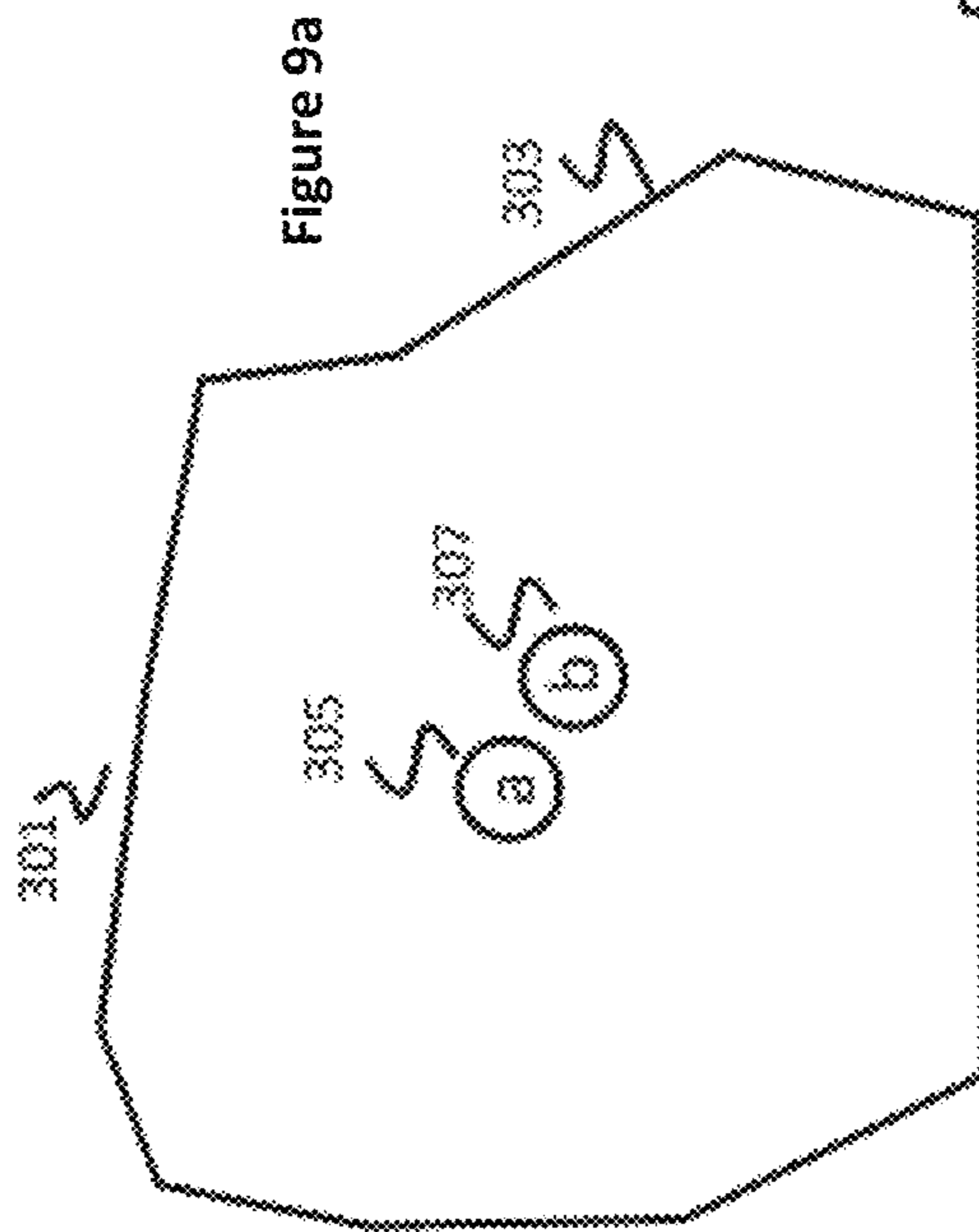


Figure 9a

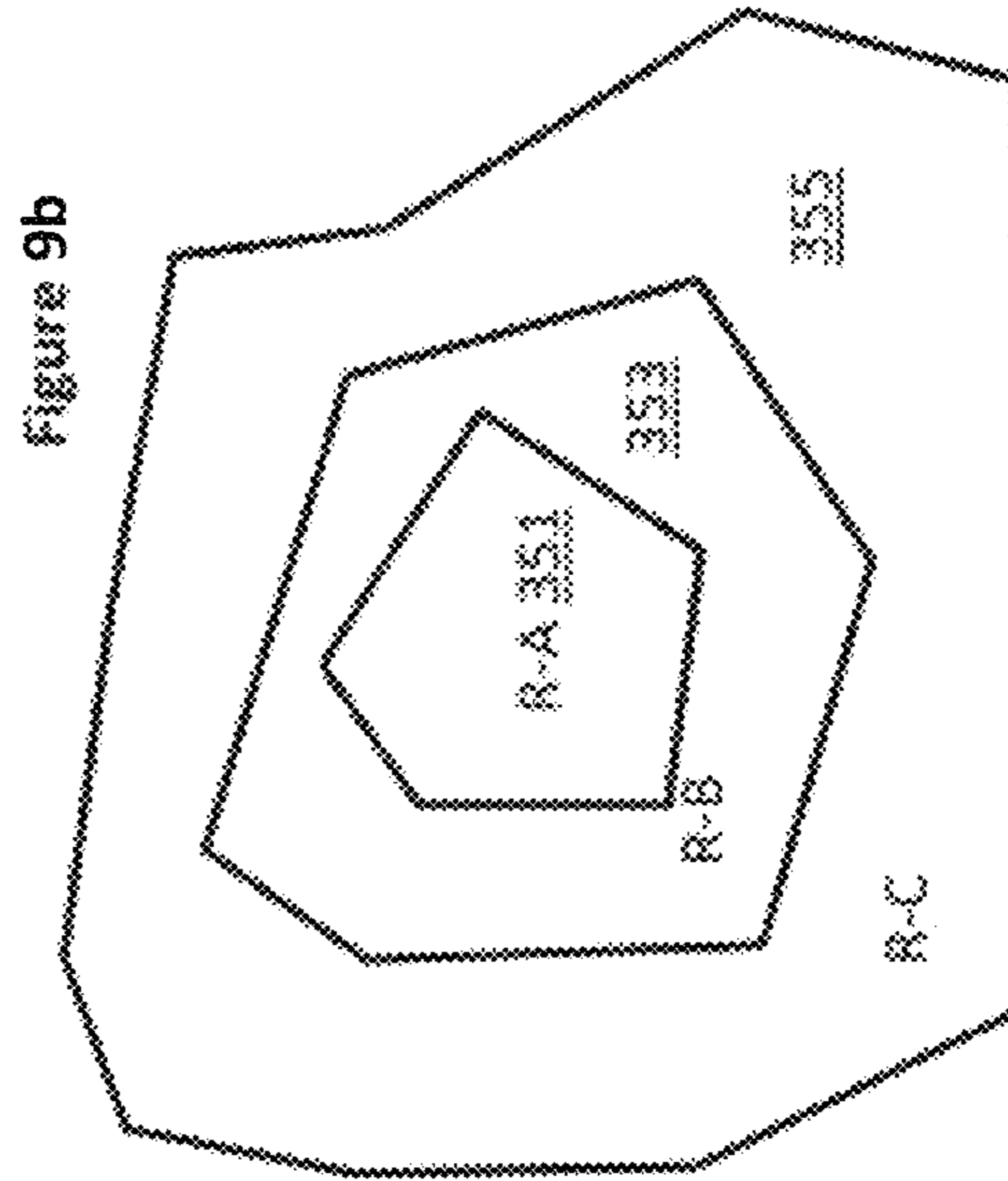


Figure 9b

Figure 10a

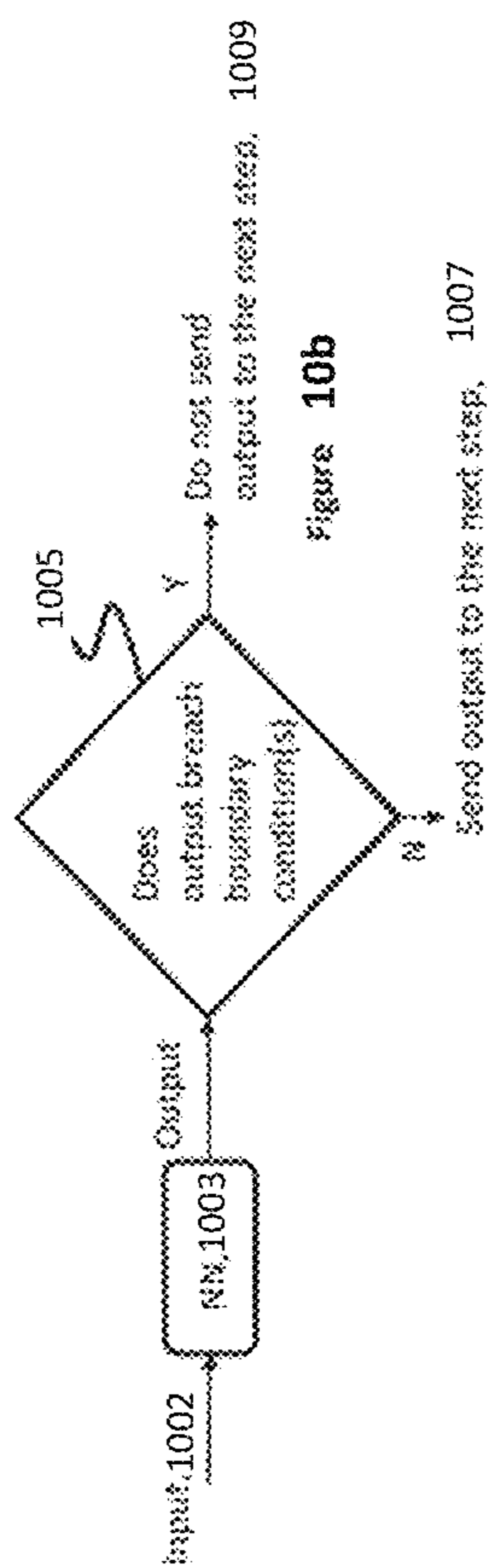
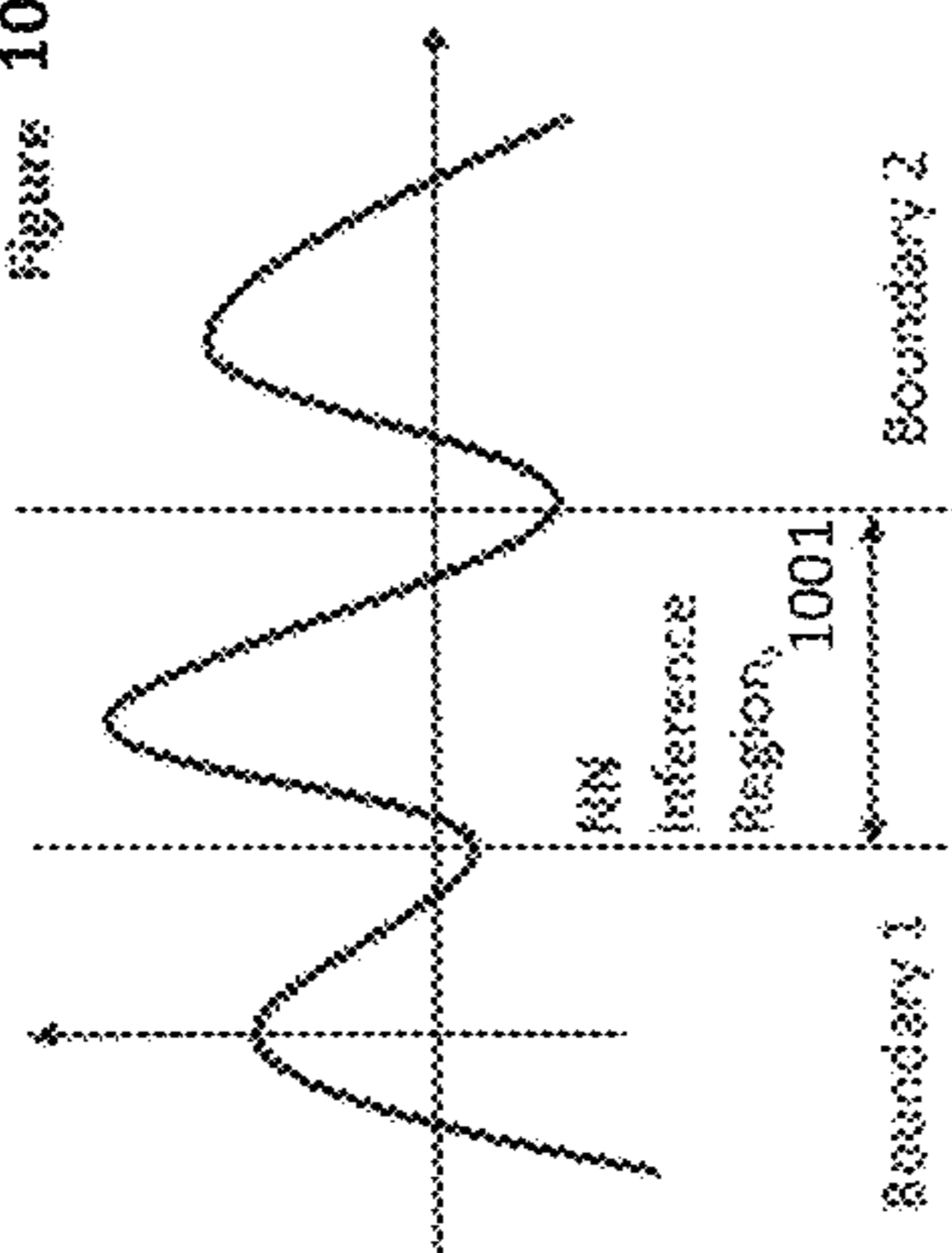


Figure 11

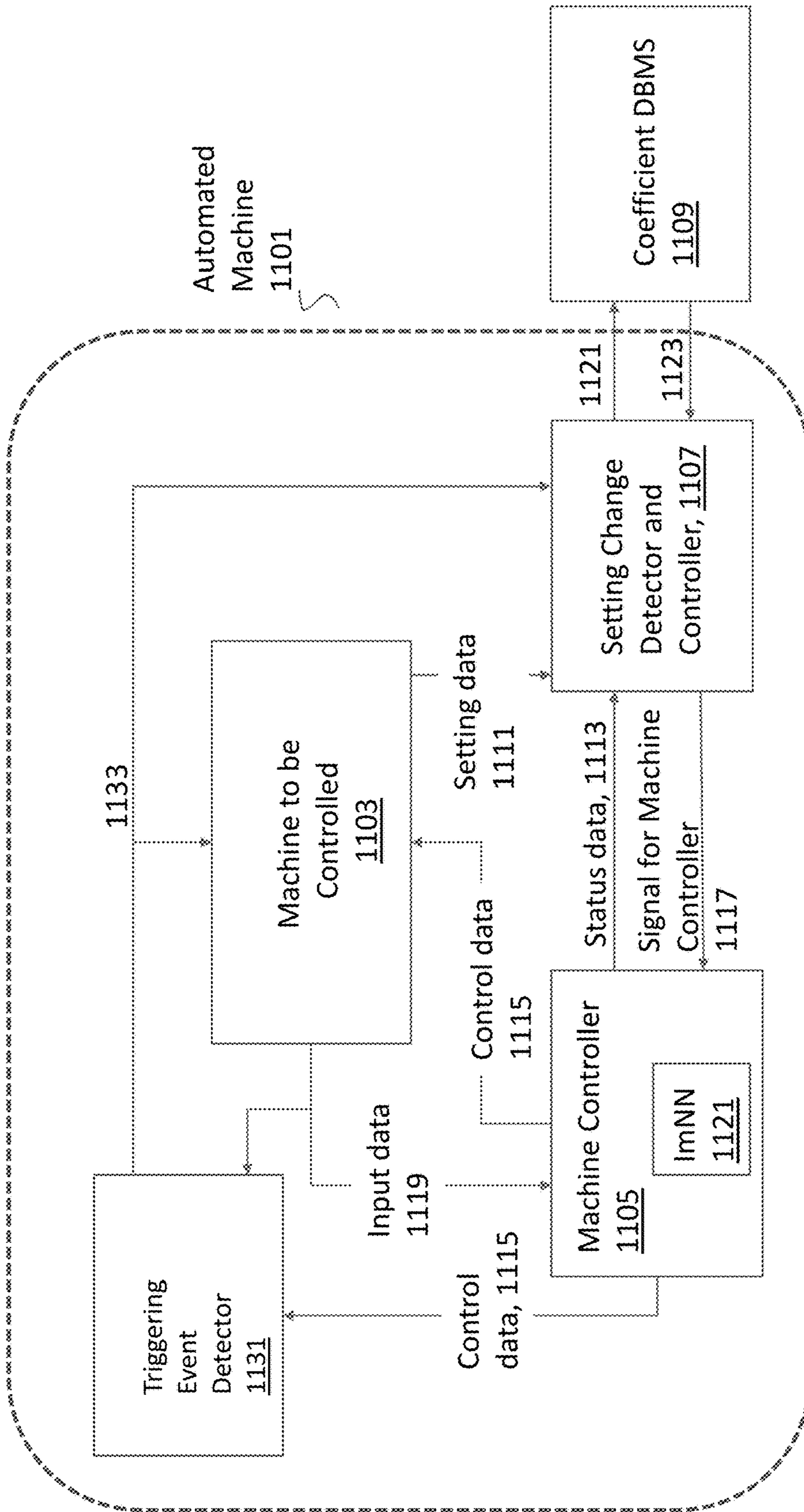
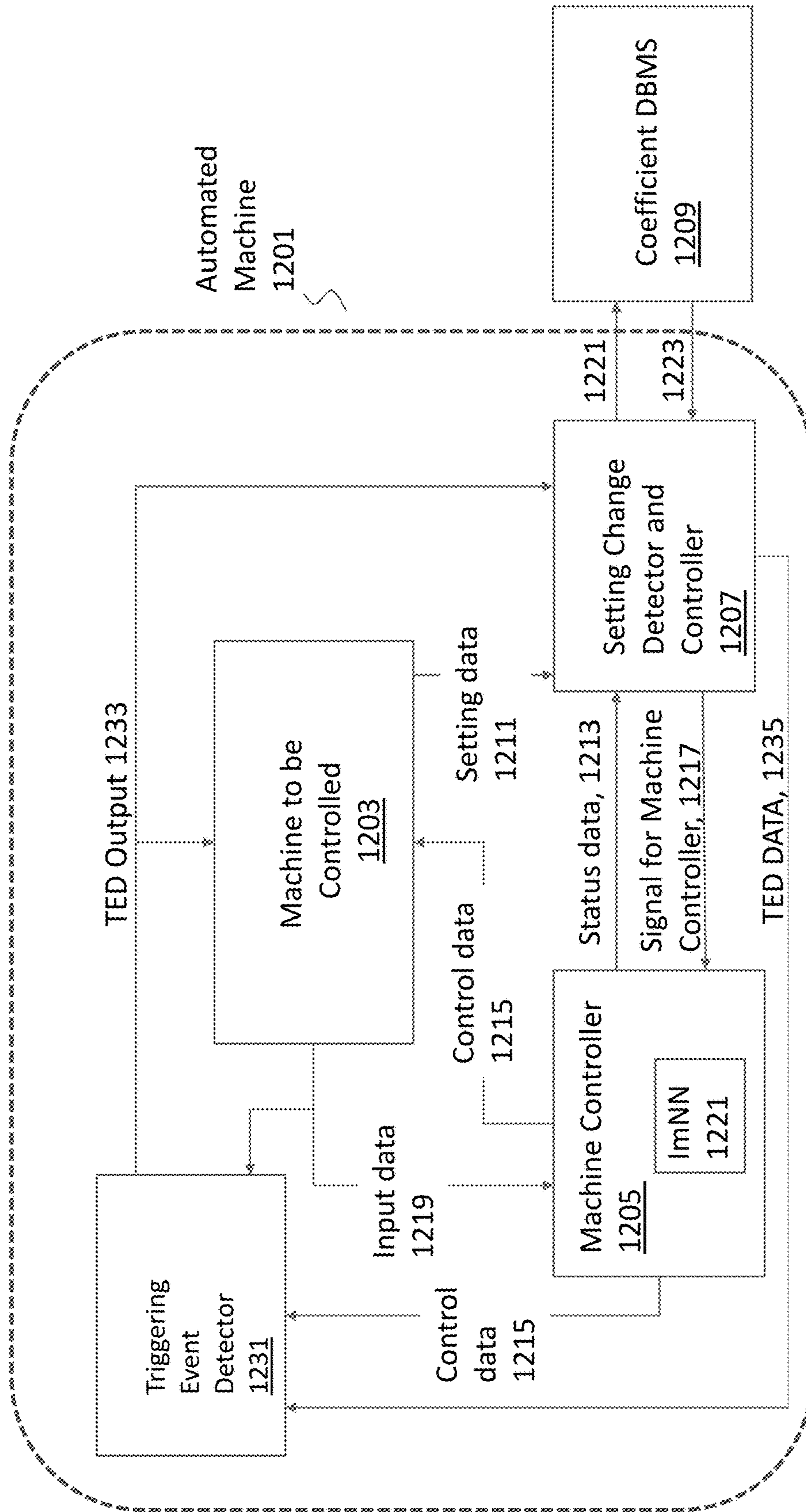


Figure 12



Automated
Machine
1201

Coefficient DBMS
1209

Setting Change
Detector and
Controller
1207

Machine to be
Controlled
1203

Machine Controller
1205
ImNN
1221

Triggering
Event
Detector
1231

TED Output 1233

Input data
1219

Control data
1215

Setting data
1211

Status data, 1213

Signal for Machine
Controller, 1217

TED DATA, 1235

1221

1223

ADAPTIVE AND INTERCHANGEABLE NEURAL NETWORKS

INCORPORATION BY REFERENCE TO ANY PRIORITY APPLICATIONS

This application claims the benefit of U.S. Provisional Application No. 62/940,762, filed on Nov. 26, 2019, and titled “ADAPTIVE AND INTERCHANGEABLE NEURAL NETWORKS.” The entire disclosure of the above-identified provisional application is hereby made part of this specification as if set forth fully herein and incorporated by reference for all purposes, for all that it contains.

Any and all applications for which a foreign or domestic priority claim is identified in the Application Data Sheet as filed with the present application are hereby incorporated by reference under 37 CFR 1.57.

BACKGROUND OF THE INVENTION

Technical Field

This invention relates to self-adapting/self-adjusting neural network system. Observing a different environment/condition/situation under which the neural network system finds itself operating from a previously observed environment/condition/situation, the system automatically reconfigures one or more neural networks—reconfiguring may include using a different set of coefficients for a neural network running within the neural network system.

Description of the Related Art

As illustrated in FIG. 1, a conventional Neural Network (NN) **101** receives an input (a single vector, in this example) at an input layer **102**, and transforms it through a series of one or more hidden layers **103**. Each hidden layer includes a set of “neurons” or “nodes,” where each neuron is connected to all neurons in the previous layer (e.g., the input layer or another hidden layer), and where neurons in a single layer function completely independently and do not share any connections. The last fully-connected layer is called the “output layer” **105**, where each of the neurons in the output layer can provide a portion of the output information (or signal), and in classification applications the output layer information (or setting(s)) represents the class scores (e.g., the score of the input(s) being classified). Based on the class scores, an output of the neural network can be determined as the class having the highest score. In other neural networks, the output may consist of an indication of the class having the highest score as the one selected and a confidence level of the selected class for given input data being the correct one based on the relative scores with other classes. This confidence level is referred to as an output confidence level, herein.

As illustrated in FIG. 2, a “convolutional” Neural Network **201** can take 3D images as input, for instance. In particular, unlike the neural network described in connection with FIG. 1, the layers of a convolutional neural network have neurons arranged in 3 dimensions: width, height, depth. Note that the word depth here refers to the third dimension of an activation volume, not to the depth of a full neural network, which can refer to the total number of layers in a network. The neurons in a layer can be configured to be connected to a small region of the layer before it, instead of all of the neurons in a fully-connected manner. The convo-

lutional neural network reduces the full image into a single vector of class scores **203**, arranged along the depth dimension.

Neural networks have been employed to operate in complex and widely varying settings (e.g., different environments, conditions, situations, and/or etc.). As such, ever increasing quantities of train data sets have been used to train prior art neural networks to operate in as many different settings as possible. When the size of training data set is increased to include training data samples from many different settings, a prior art neural network can begin to lose its accuracy and can encounter a catastrophic memory loss, which causes the neural network to cease to operate as it was originally designed and trained for.

SUMMARY OF THE INVENTION

Various aspects of the present invention includes inventive features and embodiments to allow neural network systems of the present invention to maintain or increase operational accuracy in controlling a machine while being able to operate in various different settings. In particular, a set of training data is collected over each of at least two different settings (e.g., a setting can be an environment, a condition, a situation, or the like in/under which a machine is to operate). Each setting can have its own characteristics. In some embodiments, these characteristics can be defined using a set of ranges of values. Examples of types of characteristics for the settings can be time, geographical location, and/or weather condition, etc. Using the training data set, a neural network having a particular structure can be trained for a given setting, which results in a set of coefficients for the particular neural network. For each setting, the characteristics for the setting are associated with the corresponding coefficients and/or the corresponding neural network structure trained with the training data set collected in/under the setting. Information relating to the characteristics, coefficients, and neural network structures for various settings can be stored in a database management system.

Operating in/under a setting, a neural network that has the coefficients and neural network structure associated with a set of characteristics corresponding to the setting, the neural network would yield optimal results for which it is designed/trained. In operation, various characteristics of the setting are monitored since the machine can move into or the environment/condition/situation may change to a new setting. That is, the setting may change from one setting to a new setting—the neural network coefficients and/or neural network structure (or the neural network executable module having the structure and/or the coefficients) associated with the new setting can be retrieved from a database management system. A new neural network can then be instantiated with those coefficients and may become operational, while the old neural network becomes inactive (e.g., becomes non-operational or terminated). In other words, various embodiments of the present invention allow adaptively changing the neural network(s) based on changing settings (e.g., changes in environment, condition and/or situation).

First variations of preferred methods of controlling a machine include a process (or steps of), without requiring a particular order or sequence, storing at least two sets of neural network coefficients, each being different from the others with one or more characteristics of a setting, associating each of the at least two sets of neural network coefficients with at least one set of one or more ranges of values, receiving first data from one or more input devices

of the machine, selecting one from the at least two sets of neural network coefficients based on the first data and the at least one set of one or more ranges of values. The methods of the various embodiments may also include the steps of instantiating a neural network with the selected one from the at least two sets of neural network coefficients, and controlling an aspect of the machine using an output from the instantiated neural network. As indicated above, the use of “step” herein when referring to a portion of a process does not itself indicate any particular sequence or order of the process portions, unless otherwise indicated explicitly or as required by the context of the described process.

First variations of preferred methods of controlling a machine may also include, without requiring a particular order or sequence, the steps of associating a plurality among the at least two sets of neural network coefficients with a second set of one or more ranges of values, and/or storing information relating to a neural network structure associated each of the at least two sets of neural network coefficients. The methods may further include the step of selecting one from the at least two sets of neural network further comprises the step of matching the first data with one of the at least one set of one or more ranges of values. The matching step can further comprises the steps of comparing the first data with the at least one set of one or more ranges of values; and identifying the selected one of the at least one set among one or more ranges of values that has the first data fall within its ranges of values, wherein the neural network coefficients matched with the selected one are generated by using training data set collected within the corresponding particular setting.

Second variations of preferred methods of controlling a machine may include the steps of, without requiring a particular order or sequence, storing at least two sets of neural network coefficients, each being different from the others with one or more characteristics of a setting, associating each of the at least two sets of neural network coefficients with one or more characteristics of a setting, receiving first data from one or more input devices of the machine, selecting one from the at least two sets of neural network coefficients based on the first data and the one or more characteristics of settings, instantiating a neural network with the selected one from the at least two sets of neural network coefficients, and controlling an aspect of the machine using an output from the instantiated neural network. Wherein, each of the one or more characteristics of settings is defined with a range of values.

Second variations of preferred methods of controlling a machine may also include the steps of, without requiring a particular order or sequence, storing information relating to a neural network structure associated each of the at least two sets of neural network coefficients. The step of selecting one from the at least two sets of neural network coefficients further comprises the step of matching the first data with the one or more characteristics of settings, which may further include the steps of comparing the first data with the one or more characteristics of settings, wherein each of the one or more characteristics of settings is defined with a range of values, and identifying the selected one of the one or more characteristics of settings that the first data fall within the ranges of values.

Various embodiments of preferred methods of controlling a machine may further include, without requiring a particular order or sequence, the steps of storing a set of one or more input range values associated each of the at least two sets of neural network coefficients, comparing the first data with the one or more input range values associated with the

selected one from the at least two sets of neural network coefficients, and selecting a new set among the at least two sets of neural network coefficients if the first data is outside the input range values. In other variation, the methods may include the steps of storing a set of one or more output range values associated each of the at least two sets of neural network coefficients, comparing the output with the one or more output range values associated with the selected one from the at least two sets of neural network coefficients, and selecting a new set among the at least two sets of neural network coefficients if the output is outside the output range values.

First variations of preferred apparatuses of controlling a machine may include a database management system stored with at least two sets of neural network coefficients being different from each other, at least one set of one or more ranges of values with one or more characteristics of a setting, and each of the at least two sets of neural network coefficients being associated with at least one set of one or more ranges of values; and means for controlling coupled to receive first data from one or more input devices of the machine, wherein the means for controlling includes means for selecting one from the at least two sets of neural network coefficients based on the first data and the at least one set of one or more ranges of values, and means for instantiating a neural network with the selected one from the at least two sets of neural network coefficients, wherein the neural network is configured to generate an output being used to control an aspect of the machine.

Second variations of preferred apparatuses of controlling a machine may include a database management system stored with at least two sets of neural network coefficients being different from each other with one or more characteristics of a setting, at least one set of one or more ranges of values, and each of the at least two sets of neural network coefficients being associated with at least one set of one or more ranges of values, and a controlling device that is coupled to receive first data from one or more input devices of the machine, arranged to select one from the at least two sets of neural network coefficients based on the first data and the at least one set of one or more ranges of values, and arranged to instantiate a neural network with the selected one from the at least two sets of neural network coefficients, wherein the neural network is configured to generate an output being used to control an aspect of the machine.

In the first and second variations of preferred apparatuses, the database management system may further store a plurality among the at least two sets of neural network coefficients associated with a second set of one or more ranges of values, and information relating to a neural network structure associated each of the at least two sets of neural network coefficients. In these embodiments, the database management system can be configured to match the first data with one of the at least one set of one or more ranges of values, to compare the first data with the at least one set of one or more ranges of values, and to identify the selected one of the at least one set among one or more ranges of values that has the first data fall within its ranges of values.

Third variations of preferred apparatuses of controlling a machine may include a database management system stored with at least two sets of neural network coefficients being different from each other, at least one setting having one or more characteristics with one or more characteristics of a setting, and each of the at least two sets of neural network coefficients being associated with the at least one setting having one or more characteristics, and a controlling device that is coupled to receive first data from one or more input

devices of the machine, arranged to select one from the at least two sets of neural network coefficients based on the first data and at one least one setting having one or more characteristics, and arranged to instantiate a neural network with the selected one from the at least two sets of neural network coefficients, wherein the neural network is configured to generate an output being used to control an aspect of the machine.

Fourth variations of preferred apparatuses of controlling a machine may include a database management system stored with at least two sets of neural network coefficients being different from each other with one or more characteristics of a setting, at least one setting having one or more characteristics, and each of the at least two sets of neural network coefficients being associated with the at least one setting having one or more characteristics, and means for, coupled to receive first data from one or more input devices of the machine, selecting one from the at least two sets of neural network coefficients based on the first data and at one least one setting having one or more characteristics, and instantiating a neural network with the selected one from the at least two sets of neural network coefficients, wherein the neural network is configured to generate an output being used to control an aspect of the machine.

In the third and fourth variations of preferred apparatuses, the database management system may further store each of at least one setting having one or more characteristics is defined with a range of values. The database management system also can be configured to match the first data with one of at least one setting having one or more characteristics, and may be further configured to compare the first data with the at least one setting having one or more characteristics defined with a range of values and to identify the selected one of the at least one set among one or more ranges of values that has the first data fall within its ranges of values.

In variations of preferred apparatuses of controlling a machine, the database management system can further store a set of one or more input range values associated each of the at least two sets of neural network coefficients and the instantiated neural network with the selected one from the at least two sets of neural network coefficients, and the apparatuses can further include a first trigger event detector arranged to compare the first data with the one or more input range values associated with the selected one from the at least two sets of neural network coefficients and to send a signal to the controlling device to select a new set among the at least two sets of neural network coefficients if the first data is outside the input range values. The database management system can also store a set of one or more output range values associated each of the at least two sets of neural network coefficients and the instantiated neural network with the selected one from the at least two sets of neural network coefficients, and further includes a second trigger event detector arranged to compare the output with the one or more output range values associated with the selected one from the at least two sets of neural network coefficients and to send a signal to the controlling device to select a new set among the at least two sets of neural network coefficients if the output is outside the output range values.

For the various preferred embodiments, the neural network structure can be one of a convolutional neural network, a feed forward neural network, a neural Turing machine, Hopfield neural network, and Boltzmann machine neural network. In these embodiments, a setting can be one of a temperate urban region, a desert rural region, a forested mountain region, and a coastal city and/or can be one of environment, condition, and situation in/under which the

machine operates, and the information relating to the at least two sets of neural network coefficients is stored in a standardized format to allow access by electronic devices manufactured by different manufacturers or electronic devices belonging to different manufacturing entities. Also, the neural network coefficients matched with the selected one can be generated by using training data set collected within the corresponding setting.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 is a schematic diagram illustrating a prior art neural network with hidden layers;

FIG. 2 is a schematic diagram illustrating a prior art convolutional neural network;

FIG. 3 is an illustration depicting a node on a nodal layer, the node receiving input from nodes from the previous nodal layer;

FIG. 4 is a block diagram illustrating an example preferred embodiment of an automated machine with a coefficient DBMS;

FIG. 4a is a flow chart diagram illustrating steps performed by a preferred example embodiment of a setting change detector and controller;

FIG. 4b is a timing diagram illustrating steps performed by a preferred example embodiment of a setting change detector and controller;

FIG. 4c is a timing diagram illustrating steps performed by a preferred example embodiment of a setting change detector and controller that has a predicting setting feature;

FIG. 5 is a block diagram illustrating an example preferred embodiment of a controller that can allow a plug-in architecture.

FIG. 6 is a flowchart illustrating various process steps for the system level shown in FIG. 5;

FIG. 7 is a flowchart illustrating various process steps for the neural network shown in FIG. 5;

FIG. 8 is a flowchart illustrating various process steps for the PISA module shown in FIG. 5;

FIG. 9a is a diagram illustrating a two-dimensional decision space with two potential classification groupings;

FIG. 9b is a diagram illustrating a two-dimensional decision space with multiple boundary condition regions;

FIG. 10a is a diagram illustrating a one-dimensional decision space with two potential classification groupings;

FIG. 10b is a flow chart illustrating a set of steps in using boundary conditions in a control system;

FIG. 11 is a block diagram illustrating an example preferred embodiment of an automated machine with a triggering event detector; and

FIG. 12 is a block diagram illustrating an example preferred embodiment of an automated machine with a triggering event detector with the TED signal from the SCDC.

DETAILED DESCRIPTION OF CERTAIN INVENTIVE ASPECTS

The detailed description of various exemplary embodiments below, in relation to the drawings, is intended as a description of various aspects of the various exemplary embodiments of the present invention and is not intended to represent the only aspects in which the various exemplary embodiments described herein may be practiced. The detailed description includes specific details for the purpose of providing a thorough understanding of the various exemplary embodiments of the present invention. However, it will be apparent to those skilled in the art that some aspects of the

various exemplary embodiments of the present invention may be practiced without these specific details. In some instances, well-known structures and components are shown in block diagram form in order to avoid obscuring various examples of various embodiments.

Although particular aspects various exemplary embodiments are described herein, numerous variations, combinations and permutations of these aspects fall within the scope of the disclosure. Although some benefits and advantages of certain aspects are mentioned, the scope of the disclosure is not intended to be limited to particular benefits, uses or objectives.

1. Neural Networks

Some aspects of various exemplary embodiments are described by referring to and/or using neural network(s). Various structural elements of neural network include layers (input, output, and hidden layers), nodes (or cells) for each, and connections among the nodes. Each node is connected to other nodes and has a nodal value (or a weight) and each connection can also have a weight. The initial nodal values and connections can be random or uniform. A nodal value/weight can be negative, positive, small, large, or zero after a training session with training data set. The value of each of the connection is multiplied (or other mathematical operation) by its respective connection weight. The resulting values are all added together (or other mathematical operation). A bias (e.g., nodal value) can also be added (or other mathematical operation). A bias can be a constant (often -1 or 1) or a variable. This resulting value is the value of the node when activated. Another type of nodes are convolutional nodes, which can be similar to aforementioned nodal features, are typically connected to only a few nodes from a previous layer, particularly adapted to decode spatial information in images/speech data. Deconvolutional nodes are opposite to convolutional nodes. That is, deconvolutional nodes tend to decode spatial information by being locally connected to a next layer. Other types of nodes include pooling and interpolating nodes, mean and standard deviation nodes to represent probability distributions, recurrent nodes (each with connections other nodes and a memory to store the previous value of itself), long short term memory (LSTM) nodes that may address rapid information loss occurring in recurrent nodes, and gated recurrent units nodes that are a variation of LSTM node by using two gates: update and reset.

A neural network can be a feedforward network that includes multi-level hidden layers with each layer having one or more nodes. In some exemplary embodiments of the present invention, a neural network can be a recurrent neural network either forward moving only in time or bi-directional as including forward moving components and backward moving components. Some exemplary aspects of the present invention contemplate using a recursive neural network that can configure itself adaptively with different number of layers with different number of nodes for each layer depending on given training data. In some embodiments of the present invention, the recursive neural network is a configuration of a neural network created by applying the same set of weights recursively over a structured input (producing a structured prediction over variable-size input structures) or a scalar prediction on it by traversing a given structure in topological order.

In some aspects, various exemplary embodiments contemplate taking advantage of the nonlinearity of a neural network, which may cause loss functions to become non-convex. In other words, neural networks are typically trained by using training data set on iterative, gradient-based opti-

mizers that would drive the cost function to a very low value. In some exemplary aspects of the present invention, when training data set can be preprocessed to develop characteristic by large linear regression, support vector machines with gradient descent can be used to train a neural network.

For computing the gradient (e.g., in feed-forward neural networks), in some exemplary embodiments contemplate using backpropagation, while another method such as stochastic gradient descent can be used to perform learning using this gradient. In some aspects of the present invention, the backpropagation can also be applicable to other machine learning tasks that involve computing other derivatives, e.g., part of the learning process, or to analyze the learned model.

In some exemplary embodiments, neural networks may undergo regularization (and, optionally, optimization for neural network training) during a training session using training data set. In some aspects of the present invention, regularization contemplates to be modification to the neural network to reduce its generalization error. The optimization, in some exemplary embodiments, can use continuation methods. This option can make optimization more efficient by selecting initial points causing the local optimization efforts in well-behaved regions of training data set space. In another exemplary embodiment, the optimization can use a stochastic curriculum, e.g., gradually increasing the average proportion of the more difficult examples is gradually increased, whereas in a conventional training a random mix of easy and difficult examples is presented to neural nets to be trained.

In some exemplary embodiments, supervised training or unsupervised training (or combination thereof) can be employed to train a given neural network. The unsupervised training allows a neural network to discern the input distribution/pattern on its own. In some exemplary embodiments of the unsupervised training, each layer of a neural network can be trained individually unsupervised, and then the entire network is trained to fine tune.

In some exemplary aspects of present invention, the input data are sampled so that the neural network can be more efficiently trained. In this example embodiment, sampling can be performed by using statistical methods to approximate the input distribution/pattern such as Gibbs sampling. The Gibbs sampling is an example approach in building a Markov chain, which is an example method to perform Monte Carlo estimates.

The above described various types of nodes are used in a number of different example neural network structures, such as the feedforward neural network described in connection with FIG. 1. Other example neural network structures include: a Hopfield network, a network where every neuron is connected to every other neuron; a Boltzmann machines, which is similar to the Hopfield network but with some nodes used as input/output nodes and others remain hidden nodes; and a Restricted Boltzmann machine. These three example neural network structures can include Markov chains used as preprocessors.

Another example set of neural network structures include deep convolutional neural networks and deconvolutional networks, which use the convolutional and deconvolutional nodes described above. The convolutional/deconvolutional networks can be combined with feedforward neural networks. For instance, generative adversarial networks can be formed by two different neural networks such as a combination of a feedforward neural network and convolutional neural network, with one trained to generate content related information (e.g., feature extraction) from input data and the

other trained to use the content related information to determine the content (e.g., identifying objects in images).

Another example group of neural network structures includes: recurrent neural networks that use the recurrent nodes described above, LSTM use the aforementioned LSTM nodes, gated recurrent units having an update gate instead of other gate of LSTM, neural Turing machines that have memories separated from nodes, bidirectional recurrent neural networks, and echo state networks having random connections between recurrent nodes.

Yet another example group of neural network structures includes: deep residual networks which is a deep feedforward neural networks with extra connections passing input from one layer to a later layer (often 2 to 5 layers) as well as the next layer, extreme learning machines that is a feedforward neural network with random connections but not recurrent or spiking. In some implementations the deep feedforward neural network has more than five layers. Regarding a spiking neural network, liquid state machines are similar to extreme learning machines with spiking nodes, such as replacing sigmoid activations with threshold functions and each node has a memory capable of accumulating.

Other example structures include support vector machines that finds optimal solutions for classification problems, self-organizing neural networks such as Kohonen neural networks. Another example set of neural network structures includes: autoencoders configured to automatically encode information, sparse autoencoders that encode information in more space, variational autoencoders are pre-injected with an approximated probability distribution of the input training samples, denoising autoencoders that train with the input data with noise, and deep belief networks are stacked structures of autoencoders. The deep belief networks have been shown to be effectively trainable stack by stack.

In some embodiments, the neural network may include a neural network that has a class of deep, feed-forward artificial neural networks that use a variation of multilayer perceptrons designed to require minimal preprocessing and may also use hidden layers that are convolutional layers (or CNN), pooling layers, fully/partially connected layers and normalization layers. Some embodiments can be referred to as shift invariant or space invariant artificial neural networks (SIANN), based on their shared-weights architecture and translation invariance characteristics. A neural network may self-train (e.g., Alphago Zero) such as by using re-enforcement learning. Variations on this embodiment include the deep Q-network (DQN) which is a type of deep learning model that combines a deep CNN with Q-learning, a form of reinforcement learning. Unlike earlier reinforcement learning agents, DQNs can learn directly from high-dimensional sensory inputs. Variation on this embodiment include convolutional deep belief networks (CDBN) which have structure very similar to the CNN and are trained similarly to deep belief networks. These extensions exploit the 2D structure of images, like CNNs do, and make use of pre-training like deep belief networks. Further variations on this embodiment include time delay neural networks (TDNN) which allow timed signals (e.g. speech) to be processed time-invariantly, analogous to the translation invariance offered by CNNs. The tiling of neuron outputs can cover timed stages. It should be noted that the above-mentioned neural networks can be trained using training data sets using the unsupervised learning, the supervised learning, or the reinforcement learning steps.

2. Nodal Coefficients and Database Management System for Such Coefficients

FIG. 3 illustrates an example of nodal coefficients (or nodal weights—nodal weights and nodal coefficients can be used interchangeably in the context of various embodiments of the present invention, and they can be shortened to coefficients or weights) for a node in a feed-forward neural network. In this particular example, there are “n” number of coefficients per a node **321** including a bias coefficient **323**. The output **325**, a_{out} , is an output of a function applied to the sum **329** (Σ) of (i) a bias value ($1*b$) and (ii) each output from the previous layer of nodes (a_1, a_2, \dots, a_N in FIG. 3) with each multiplied by a certain weight (W_1, W_2, \dots, W_N in FIG. 3) that can be a negative/positive number or zero. The function **327** (e.g., $z \rightarrow g$) applied to the sum can be predefined. Using a set of equations, these can be expressed as:

$$Z = b + \sum_{i=1}^N a_i w_i$$

$$a_{out} = g(z)$$

The weights are determined by training the given neural network with a training data set, which can include multiple input and output data pairs. In some preferred embodiments, training data set is for self-learning neural networks. Although the preferred embodiment in FIG. 3 depicts a feedforward neural network structure that sums the bias and each output from the previous layer multiplied by a coefficient, and then applying a function to the sum, other preferred embodiments contemplate the use of various i) neural network structures (examples described in Section 1 above), ii) arrangements regarding which nodes on which previous layers (e.g., not just immediately previous layer but a sub-set or all previous layers) would be connected, and/or iii) other operation(s) can be used (e.g., multiplication operation instead of or in addition to the summing operation).

As noted above, prior art neural networks have been trained and/or employed to operate in complex and widely varying settings. A “setting” as used herein, refers generally to any particular environment or location, such as a particular condition, environment, situation and/or etc. In various embodiments of the present invention, rather than or in addition to training a neural network with ever increasing quantity of training data sets to cover various settings, a neural network coefficient set for a neural network structure is trained using a training data set collected for a particular setting. Non-exhaustive various example types of characteristics of various settings may include:

For vehicles, drones, missiles, etc.:

Geographical environments: high population density city, medium population density city, suburbia areas, rural regions, plain regions, mountainous areas, coastal areas, etc.

Weather conditions—raining, snowing, clear day, windy day, foggy, etc.

Situations—amount of surrounding traffic, accident ahead, wild animals present, etc.

For speech recognition machines:

Geographical: Midwestern, southern, north eastern, etc. in the US

Recognizable accents: Midwestern, Northeastern, Southern, etc. in the US

For facial recognition machines:

Geographical: continents, countries, cities, rural areas, etc.

Ethnic background: Northern African, Sub-Saharan African, Northern European, Southern European, Eastern European, Northeast Asian, Southeast Asian, Middle Eastern, etc.

For an assembly line application:

Type of objects to be sorted or type of operations to be performed on objects

For target identification in military applications:

Different targets and/or in different environments or situations, some examples are: a. for targeting tanks/artillery pieces/missile launchers on the surface of desert v. in forested regions from aircraft/drones, b. for targeting drones/aircraft on a sunny day, full moon/no moon night, or rainy day/night from the surface

Characteristics of a setting can also relate to conditions of the sensors that generate input to neural networks, as examples: the age of sensors; the manufacturer of sensors; or the different products lines/periods even from the same manufacturer. That is, new and many-years old sensors from the same manufacturer and same production line may give rise to having to use two different sets of coefficients and/or neural network structure. In some embodiments, a setting can include a sensor being in a non-working condition. In these example embodiments, one set of coefficients with a neural network structure can be trained under/in the setting that full collection of sensors/devices functioning optimally, and other sets of coefficients with different neural network structures trained under/in the settings when one or more sensors/devices are malfunctioning.

In various embodiments of the present invention, a training data set can be separately collected from each setting in/under which the automated machine is to operate. For example, one training data set can consist of data collected using various sensors in a setting that can be characterized as a desert area, country side, and during day time with no wild animal activities. Another setting for a training data set can be characterized as data collected from a desert area, a suburban region, and during night time with some wild animal activities. A different setting for another training data set can consist of data collected using various sensors in a large city environment, during night time, and with a large number of pedestrians. For each of these different training set, a particular neural network—setup with a neural network structure—is trained. A trained neural network results in a set of coefficients (that is each node ends up with coefficients after a training session) for the particular setting in/under which training data set is collected. It should be noted that in some preferred embodiments, in addition to having a training data set for each setting, similar settings can have the same set of coefficients and neural network structure (e.g., downtown New York City and downtown Boston can use the same set of coefficients, sand dunes in the Sahara Desert and sand dunes in the Death Valley can use the same set of coefficients, etc.).

In an exemplary preferred embodiment to describe a coefficient set, a neural network can be a feedforward network and can have an input layer (e.g., five nodes), an output layer (e.g., three nodes), and five hidden layers with five nodes each. In this structure, the example neural network has 25 nodes among the hidden layers. Using the nodal coefficient example depicted in FIG. 3, each hidden layer node and each output layer node can have five weights/coefficients including a bias weight value in the example neural network. Once a set of coefficients are determined from training using the data set collected in a particular setting, such a set is associated with characteristics to the setting (as exemplified above: desert, daytime, countryside,

etc.) as described below in connection with Table 1 (and also may further described in connection with Tables 2 and 3). In this preferred exemplary embodiment, all stored coefficient sets are for a feedforward network having an input layer (five nodes), an output layer (three nodes), and five hidden layers with five nodes each. To store information/data relating to a large number of different settings, a database management system can be employed. By segmenting settings, each setting may exhibit a constant/consistent environment, condition, and/or situation, and a neural network trained for a specific setting may become more accurate within the trained setting.

It should be noted that in some other preferred embodiments of the present invention, the database can include coefficient sets with different neural network structures depending on the optimal structure for different settings. For instance, a number of sets of coefficients can be for feedforward networks, while other sets can be for back propagation networks or other neural network structures such as those provided above in Section 1, for example.

In some preferred embodiments, in addition to the coefficients, the database can store:

More information about the neural network structure such as the type of structure, the number of layers, the number of nodes on each layer, nodal connection information between and/or within each layer, and/or other information to define a neural network; and/or Executable modules of neural networks having the neural network structures and corresponding coefficients.

The automated machine can broadly refer to a machine that is to be controlled by a control mechanism, with some human intervention if necessary. Examples of an automated machine can be appliances (e.g., ovens, refrigerators) with automated controllers (e.g., Internet of Things, “IoT” controllers), a speech generator, a speech recognition system, a facial recognition system, an automated personal assistant (e.g., Alexa by Amazon, Inc., Siri by Apple, Inc.), an autonomous vehicle, a robot, a target recognition system (e.g., in military defense applications such as missiles and drones), and etc. Also, an automated machine does not necessarily mean a completely automated manual-less machine that requires no human intervention, but it may require a qualified person to take over the control (e.g., driving) under certain circumstances.

In one example preferred embodiment illustrated in FIG. 4, automated machine 401 may include a machine-to-be-controlled (“MTBC”) 403, a machine controller (“MC”) 405, and a setting change detector and controller (“SCDC”) 407. The system illustrated in FIG. 4 also includes a coefficient database management system (“C-DBMS”) 409, which is depicted as located outside the automated machine 401 while operationally coupled thereto. Note that in such an embodiment, the C-DBMS 409 can be located at a remote server and coupled to communicate with the automated machine. In some other preferred embodiments, a sub-set or the whole of the C-DBMS 409 can be co-located with an automated machine 401 or can be considered as a part of the automated machine. In yet another embodiment, the whole of the C-DBMS 409 can be co-located with the automated machine 401 and implemented on hardware or firmware for fast accesses of the information stored therein.

In FIG. 4, the MTBC 403, MC 405, SCDC 407 and C-DBMS 409 are described as individual modules. As such, they can be located remote from each other. For instance, in an example windmill embodiment, the machine can be the components of a wind-turbine, the MTBC 403 can be located proximate to or on the wind-turbine to receive input

from input devices (e.g., wind speed sensor) and to control the speed of the generator, the angle of the blades, the rotation of the wind-turbine. Continuing with the windmill example, the MC 405 and the SCDC 407 can be located at the base of the windmill for the ease of access, and the C-DBMS 409 can be located at a server located remote from the windmill. It should be noted in some preferred embodiments, the MTBC 403, MC 405, SCDC 407 and C-DBMS 409 are one whole module rather than individual modules. It should also be noted that, although the MTBC 403 is named as “machine to be controlled,” MTBC 403 preferably may include various interfaces to receive data from various sensors and/or devices and various interfaces to send control information to components (e.g., motors, actuators, loud speakers) in controlling the automated machine on which the MTBC 403 is to control.

It should be noted that the MTBC 403, MC 405, SCDC 407 and C-DBMS 409 can be implemented on/as hardware, firmware, software modules or combination of them. In case of being software modules, those modules can be implemented as virtual machine(s) and/or software container(s).

Input data 419 generated by the MTBC 403 is sent over to the MC 405 to be processed (e.g., inferenced) by an ImNN 421. The MC 405 generates control data 415, which is sent over to the MTBC 403. The MC 405 also generates status data 413 for the SCDC 407, and the SCDC 407 uses a signal for the MC 417 to control the life cycle of the ImNN 421 (e.g., instantiate, terminate, run, and etc.) The values of setting data 411 are sent from the MTBC 403 to the SCDC 407. In various embodiments of the present invention, the values of setting data 411 (although “of” is used, in various embodiments of the present invention, the values can be also described as “on” setting data 411 as on a data bus or obtained “from” the setting data 411 as in from shared memory) can be seen as data collected/captured/sensed by various of sensors relating to the setting (e.g., environment, condition, situation, and etc.). The values of the setting data 411 can be referred to as setting characteristic values.

In FIG. 4, the MTBC 403 is coupled with the MC 405 via the input data 419 and control data 415, the MC 405 and the SCDC 407 are coupled with the status data 413 and signal for MC 417, the MTBC 403 and the SCDC 407 are coupled with the setting data 411, and the SCDC 407 and C-DBMS 409 are coupled with two directional arrows 421 and 423—the coupling 421 sends queries from SCDC 407 to the C-DBMS 409 and the coupling 423 sends results of the queries to the SCDC 407. These above-mentioned coupling mechanisms, and those illustrated in FIGS. 11 and 12 below, provide communication mechanisms to send and/or receive data and/or control signal(s). The coupling mechanism can be implemented using, for example, shared memory, sockets (in socket communication) and/or hardware implemented data/control signal buses. In particular, the input data 419 sends data from various sensors/devices located in the MTBC 403 to the MC 405, control data 415 sends data to control various controllable components interfacing with the MTBC 405, the status data 413 sends data from the MC 405 about the status of the ImNN(s) 412, the signal for MC 417 is the data to control the ImNN(s), and the setting data 411 sends data from various sensors/devices in the MTBC 403—that is, the setting characteristic values.

With respect to the MTBC 403, it includes (or has interfaces to) various input sensors/devices, communication devices and machine control devices, such as a thermometer, pressure sensor, compass, altimeter, gyroscope, accelerometer, image sensor, cameras, video cameras, magnetometer, light detectors (e.g., visible, infra-red, ultra-violet), barom-

eter, humidity measuring device, radiation sensor, audio/sound sensor, e.g., microphone, geographical positions system (GPS) device, ground to surface distance (GSD) device and/or etc. From these input sensors/devices various setting characteristic values can be obtained. For example, the temperature from a thermometer, air pressure (e.g., of a tire) from a pressure sensor, magnetic North from a compass, altitude from an altimeter, orientation information from a gyroscope, acceleration information from an accelerometer, images from an image sensor or camera, video frames from a video cameras, magnetic field information from a magnetometer, ambient light variation information from light detectors, atmospheric/ambient air pressure from a barometer, humidity level from humidity measuring device, radiation level from a radiation sensor, voice from audio/sound sensor, geospatial information from a GPS device.

In an autonomous land vehicle example, the MTBC 403 may include (or interfaces to) a number of sensors and internal computing devices, with the following examples, to control the vehicle while traveling in traffic with other land vehicles. Sensors for collecting external surrounding information include one more front view cameras (e.g., digital camera), a night vision camera(s), a front object laser sensor(s), front and rear millimeter radars and sensors, an ambient light sensor, pedestrian/animal detecting IR sensor (s), a side view camera(s) on each side, a night vision camera(s) on each side, a proximity sensor(s) on each side, a panoramic/wide angle view sensor(s) (e.g., 100 degrees, 180 degrees, and/or 360 degrees view digital cameras), a LIDAR sensor, a tire pressure sensor for each mounted tire, a wheel speed sensor for each wheel, a rear view camera(s) (e.g., digital camera), and/or a review view night vision camera(s). As used herein, a “camera” is a broad term and refers to any of a number of imaging devices/systems that collect data representative of an “image” (e.g., a one or multi-dimensional representation of information) with one or more sensors (e.g., film or one or more electronic sensors), unless the context of the usage indicates otherwise. The number of cameras and sensors having various views may be mounted on an autonomous land vehicle so that, preferably, there are no gaps or blind spots either going forward or backward. Sensors can also include GPS devices, gyroscopes, and etc. that give the direction, velocity, and/or location information of the automated machine.

Moreover, sensors for collecting operational information and having interfaces with the MTBC 403 include a driver drowsiness sensor, steering angle sensor, a throttle (e.g., gas pedal) pressure sensor, and/or a bread pedal sensor. In addition to sensors, the autonomous vehicle may also include communication devices to send and receive data from a network (e.g., cell phone network, Wi-Fi, GPS and/or other types of communication networks that provide secured communication method) and from other vehicles via vehicle-to-vehicle communication networks (e.g., VANETs) that provides secured communication links. These devices may also interface with the MTBC 403.

The autonomous vehicle may be configured to include or to interface with a communication device (e.g., a cell phone, radio, or the like) on its own within to interface with the MTBC 403 or include a docking system to connect to a communication device. If the autonomous vehicle includes a docking system to connect to a cell phone and has no other means of connecting to the cell phone network, such a vehicle may provide an additional anti-theft feature by disabling the automated driving function or disabling the entire driving function without being connecting to the communication network with the communication device.

Machine control devices interfacing with the MTBC 403 for the autonomous land vehicle may include (or include interfaces to) adaptive cruise control, an on-board computer(s), one or more control chips and/or control mechanisms to control the breaking, throttle, and steering wheel systems. Machine control devices interfacing with the MTBC 403 for a drone having fixed wings may include mechanisms to control elevator(s), flap(s), and/or aileron(s), in addition mechanisms to control the thrust(s) and the ruder. If a drone has rotor(s), the MTBC 403 may include (or has interfaces to) a control mechanism for the rotor(s). Machine control devices within the MTBC 403 for a missile with aerodynamic devices (e.g., canard(s), wing(s), and/or tail(s)), may include (or has interfaces to) control mechanisms for those devices. Machine control devices within the MTBC 403 for a robot may include (or has interfaces to) control mechanisms for various actuators (e.g., pneumatic actuators, hydraulic actuators, and/or electric actuators.) For a speech generator, a control mechanism (or interface thereto) may control input to loud speakers. Automated machines such as drones, missiles, robots or the like can include various types sensors/devices or interfaces thereto as described above for particular use of those machines. It should also be noted that a cell phone can be an automated machine as used herein since a cell phone can have sensors (e.g., microphone(s), camera(s)) to generate input to a facial recognition system, a finger print recognition system, a speech recognition system, or a speech generator.

Continuing on with FIG. 4, the MC 405 can include one or more implementation neural network (ImNN). At the initial stage, in a preferred embodiment, the SCDC 407 can instantiate an ImNN with a default neural network structure with a default set of coefficients—e.g., a fork ImNN process is created with a default set of coefficients with a default neural network structure, via the signal for the MC 405, as shown in FIG. 4. In this embodiment, the automated machine can start its operation with the default arrangements. As the SCDC 407 receives the setting characteristic values from the MTBC 403, it determines whether to keep the default set or to query the C-DBMS 409, which is described in more detail in connection with Table 1 (and also may be further described in connection with Tables 2 and 3).

In another preferred embodiments, as illustrated in FIG. 4a, control data 415 is not sent to MTBC 403 or input data 419 is not inferred by MC 405 until:

a. (step 451) receives values of setting data 411 relating to the setting characteristics from the MTBC 403,

b. (step 452) queries the C-DBMS 409 with the received setting characteristic values which (step 453) returns A) a set of coefficients and/or a neural network structure associated with the received setting characteristic values or B) the neural network executable module (or a pointer thereto) having the structure and/or the coefficients—based on the information stored in the C-DBMS 409,

c. (step 454) instantiates a new ImNN using A) the set of coefficients and/or the neural network structure or B) the neural network executable module, and

d. the new ImNN becomes operational.

Various sensors/devices on the MTBC 403 can generate input data to be sent to the MC 405, which in turn use the input data to generate control data after conducting inferences on the input data. Here, all or a subset of input data can be inferred on by the ImNN utilizing the set of coefficients and the neural network structure used in instantiating the ImNN.

Some of the sensors/devices on the MTBC 403 may generate the setting characteristic values for the SCDC 407.

These sensors/devices can be the same sensors/devices, a subset of sensors/devices, or a different set of sensors/devices (that may include a subset of sensors/devices) on the MTBC 403 or elsewhere on the automated machine that generates input data. The SCDC 407 can continually or periodically (e.g., every fraction of a second, a second, a minute, or etc.) receive the setting characteristic values—individually, sub-set at a time, or all at once with/without a notice signal (e.g., an interrupt signal)—from the MTBC 403. The notice signal notifies the SCDC 407 that a set of setting characteristic values are prepared and will follow.

Subsequent to the ImNN becoming operational (“the currently operational ImNN”), the automated machine may move into or may be encountering a different geographical region, environment, or situation (e.g., the time of the day, weather, etc.). The information relating to the environment, condition, situation, and/or etc. (i.e., setting characteristic values) is received by the SCDC 407 as noted above. If a change in the setting is sensed (e.g., day turns to evening, sunny to cloudy, country side environment to suburban environment), the C-DBMS 409 is queried, using the current set of the setting characteristic values.

More specifically, in some embodiments of the present invention, the SCDC 407 may determine to query the C-DBMS 409 based on one or more sensor/device data. For example, the SCDC 407 can be prearranged such that when weather changes from warm to cold (e.g., with specific temperature threshold), the C-DBMS 409 is queried using the current set of setting characteristic values received from the MTBC 403. In another example, when output of a clock indicates a sunset time according to the seasonal and geographical location information, the C-DBMS 409 is queried using the current set of setting characteristic values received from the MTBC 403. In another example, the output from a light sensor can be used to cause the SCDC 407 to query C-DBMS 409 using the current set of setting characteristic values received from the MTBC 403. In some other preferred embodiments, the SCDC 407 can determine to query the C-DBMS 409 periodically (e.g., every minute, certain number of minutes, tens of minutes, etc.) using the current set of setting characteristic values received from the MTBC 403. In yet some other embodiments, the SCDC 407 can determine to query the C-DBMS after elapse of a certain amount of time since the last query to the C-DBMS 409 using the current set of setting characteristic values received from the MTBC 403. In some other embodiments, the SCDC 407 can determine to query each time a set of setting characteristic values are received from the MTBC with the notice signal using the current set of setting characteristic values received from the MTBC 403. Various events described above that causes querying the C-DBMS 409 can be used individually or a combination thereof.

It should be noted that, after instantiating a new ImNN and having it process through input data to start generating output may take a number of clock cycles—a transition phase. The currently operational ImNN can be designated as a to-be-terminated ImNN during the transition phase. In some embodiments, as illustrated in FIG. 4b, during the transition phase, the to-be-terminated ImNN (that can be referred to as the old ImNN) can continue to run until a new ImNN is properly initiated (e.g., another fork created and instantiated with the coefficient set) and becomes operational (e.g., receiving input data and generating output data). In these embodiments, subsequent to (or simultaneous with) the new ImNN becoming operational, the to-be-terminated ImNN can be terminated. In another embodiment, during the transition phase, the to-be-terminated ImNN can be termi-

nated at the end of the phase but run slowly (e.g., generating output every other clock cycle)—in this example, the to-be-terminated ImNN may not generate optimal output during the transition phase.

In another example embodiments, the next setting may be predicted. That is, as a vehicle moves from a country side towards a city, the SCDC 407 can be configured to predict the approaching city setting (e.g., by calculating the speed, the direction, and GPS information) and can be configured to instantiate a new ImNN with the city characteristics before the actual arrival at the city (e.g., with A) a new neural network executable module or B) the coefficients and/or structure, queried from the C-DBMS 409—that is, queried with predicted setting characteristic values). In these example embodiments, the new ImNN can start inferencing the input data and generating output at or before the vehicle crosses the city boundary from the country side. In other words, the new ImNN may run simultaneously with the current ImNN, but the output from the current ImNN may be used to control such a vehicle, as illustrated in FIG. 4c. As the vehicle crosses from a country side into a city environment, the output from the new ImNN may be used to control the vehicle while the current ImNN (that can be referred to as the old ImNN) is terminated. The new ImNN may run until a new setting is detected or predicted. In these embodiments, the length of the transition phase can be shortened. In FIGS. 4b and 4c, time flows from the left side to the right side.

Similar embodiments to shorten the transition phase can be contemplated with, for example, changing time (e.g., predicting the day time changing to evening time or night time changing to morning time), weather (e.g., approaching storm), temperature (e.g., from weather forecast), traffic congestion (e.g., from traffic report), and etc. It should also be noted that if the approaching setting is not predictable with certainty (e.g., weather forecast), multiple ImNNs can be pre-instantiated (e.g., based on possible approaching weather patterns).

Although FIG. 4 is explicitly showing one ImNN instantiated as part of the MC 405, in various embodiments of the present invention, more than one ImNN can be instantiated and become operational. In some preferred embodiments, various ImNNs (that is, they may have same/different sets of coefficients and/or same/different neural network structures) can be connected serially such that output from one ImNN are further inferenced by another ImNN. In some other preferred embodiments, various ImNNs can be connected in parallel such that input from one input source can be inferenced by various multiple ImNNs. Intermittent between various ImNNs, there can be logic/algorithm inserted to further process data, in some preferred embodiments (e.g., adding, combining, multiplying, and thresholding outputs from the various ImNNs). In the case of multiple ImNNs, the SCDC 407 has the corresponding control mechanisms for each of the instantiated ImNN.

The C-DBMS 409 can include searchable information associated with each setting. That is, for each setting, the C-DBMS 409 can include information on ranges of setting characteristic values (which can also be referred to as setting characteristic value ranges) and an associated set of coefficients and/or neural network structure. The C-DBMS 409 can be searched based on the setting characteristic values to find a set of coefficients and/or neural network structure for a given setting characteristic values.

Table 1 below illustrates a table of searchable entries for the purpose of illustrating information that can be stored and organized into a database, such as the C-DBMS 409. Various

embodiments of the present invention contemplate using one or more of the database types: text based, document based, hierarchical, relational, or object-oriented database management systems. Also, Table 1 illustrates one-to-one relationship between the sets of setting characteristic value ranges and sets of coefficients/neural network structures. Each entry is numbered as #1, #2, #3, . . . , #n. Various embodiments of the present invention allow many-to-one or one-to-many relationships between the set of setting characteristic value ranges and the set of coefficients/neural network structures.

TABLE 1

Entry #	Time	Location	Weather	Ambient Temp	Structure	Coefficient Array
1	Day time range	1 st Ranges of Latitudes Longitudes	Sunny	Above freezing	feedforward	1 st set of coefficients [. . .]
2	Day time range	2 nd Ranges of Lats + Longs	Cloudy	Above freezing	feedforward	2 nd set of coefficients [. . .]
3	Night time	3 rd Ranges of Lats + Longs	Rain	Above freezing	Back propagation	3 rd set of coefficients [. . .]
. . .						
n	Evening twilight	N th Ranges of Lats + Longs	Sunny	Freezing	Restricted Boltzmann	N th set of coefficients [. . .]

Although Table 1 illustrates various pieces of information (e.g., setting characteristic value ranges, coefficients, and structure) that are placed in one location (that is, Table 1), various embodiments of the present invention contemplated other embodiments in which the pieces of information can be located in remote locations from each other but linked for the database to function.

Table 1 above depicts various information that can be stored in C-DBMS 409. The top row lists example setting characteristic types: time, location, weather, ambient temperature. The top row also lists descriptive names for other columns: structure and coefficient array. The top row is provided for the ease of explaining various columns of information. In this example, the Time refers to input from a clock, the location refers to latitude and longitude from a GPS device, the weather refers to information from a barometer, a light sensor and/or a moisture sensor, and the ambient temperature refers to input from a thermometer.

The descriptive names “Structure” refers to a neural network structure, and “Coefficient Array” refers to a set of nodal coefficients for the neural network structure. In various preferred embodiments, the information contained in the columns of Coefficients Array and Structure combined is sufficient to instantiate corresponding neural network(s) for the associated setting.

In first example preferred embodiments, each of the entries has a neural network structure, which may include a pointer to an executable module in a library of compiled sets of executable modules of neural networks. For instance, a library for Table 1 could include pointers to the executable modules of neural networks for feedforward, back propagation, and Restrict Boltzmann types (although other types can

be also included). For a specific example, using entry #1, the “feedforward” in the column designated as the Structure can be a pointer to a particular version of a feedforward neural network executable module trained with training data set from the associated setting. The SCDC 407 can use the executable module and the set of corresponding coefficients in entry #1 to instantiate the feedforward neural network.

In second example preferred embodiments, the executable neural network modules may could already have been compiled with a specific set of nodal coefficients. For these example executables, the column in Table 1 designated as Coefficients Array may not be necessary—the pointers to the associated neural network executable module may be sufficient to instantiate the specified neural network, since these modules already have the coefficients compiled therein. Although the first and second examples of preferred embodiments above have been described in terms of computer programs/libraries, the library of neural networks can be implemented in hardware, firmware or combinations of hardware, firmware and software modules. In addition, instead of pointers, the modules themselves can be stored on the database as entries.

In various other preferred embodiments, the entries for the Structure entry may include information relating to the type of neural network and its basic layout, for example, nodal layers—input, output, hidden—and types of nodes, such as input node, hidden node, memory node, different memory node, convolutional node, probabilistic node, and etc. sufficient to generate automatically the corresponding executable neural network module—which then can be instantiated with the corresponding set of coefficients. In some of such preferred embodiments, the generated executable module then can be stored in the C-DBMS 409 for later use. It should also be noted that some executable neural network modules can be compiled with their corresponding coefficients, while other executable neural network modules can be compiled without coefficients already specified (for these embodiments, the entries in the “Coefficients Array” may be needed. It should be noted that a database (e.g., the C-DBMS 409) can be configured to store a mixture of entries that have pointers to neural network modules with/without coefficients already compiled therein, neural network modules rather than pointers, or information sufficient to generate executable neural network modules.

Returning back to Table 1, for each of numbered entries, ranges of values are provided for each setting characteristic type. For example, in some implementations the time can have a range indicated in hour, days, weeks, or months, or a range of a certain portion of a day (e.g., day time or night time). In some implementations, the location can have a geographic range, for example, latitudes and/or longitudes to indicate a particular region (e.g., a desert area bound by a set of latitudes and/or longitudes that can be compared with GPS data from the MTBC 403). In some implementations, the location can have other types of ranges, for example, a designated or predetermined (or predefined) area, for example, the northern hemisphere or the southern hemisphere, a country, a continent, or a region of the world. In other words, for each entry (e.g., an entry representing a setting) each type of characteristics (e.g., Time, Location, Weather, and Temperature) of a setting is defined with a range of setting characteristic values, which can be referred to as a range of values.

Various sensors/devices on the MTBC 403 may generate setting characteristic values which are matched with each entry—determining if the values received on setting data 411 fall within the ranges provided. For example, the setting

characteristic values can be: a clock indicates 10 AM, a GPS may input latitudes and longitudes that fall within the 1st ranges, a light detector may indicate sunny, and a thermometer inputs 10 degrees Celsius. In this example the received values on setting data 411 fall within the setting characteristic value ranges of the first entry. In this case, the associated set of neural network coefficients is the 1st set of coefficients and the associated neural network structure is a feedforward neural network of the entry designated as #1. Another set of setting characteristic values may match with one of ranges defined for entries #2, #3, . . . , #n.

In sum, Table 1 can be described as each entry (e.g., #1, #2, #3, . . . , #n) having setting characteristic value ranges that corresponds to characteristics of a setting. For example, if the 1st Ranges of Latitudes Longitudes may cover a desert area boundaries, this means the characteristics of entry #1 can be a setting that is a desert area, daytime, above freezing and sunny.

Although the selection process is described above as using the setting characteristic values and the setting characteristic value ranges, in other various embodiments of the present invention, the selection process can be performed by probabilistic algorithms. That is rather than search only for the entry that the setting characteristic values fall within the setting characteristic value ranges, proximity to those ranges can be calculated. The entry being the closest (e.g., having the largest number of the setting characteristic values fall within the given setting characteristic value ranges) to the setting characteristic values can then be selected.

Even though, the setting characteristic values are defined using numerical ranges of values in Table 1, in other preferred embodiments, other methods can be used to represent ranges. For example, in some embodiments image(s) can be used to represent the ranges (e.g., images of grey sky to represent the ranges of cloudy sky). In this example, the images representing the ranges can be further processed to turn them into a set of numerical values or use them as images in matching images received from a camera.

In some embodiments of the present invention, a subset of the setting characteristics can be used to locate the coefficient arrays. In another preferred embodiment, more types of setting characteristic values from different sensors/devices can be added as indicated by the column with “. . .” (e.g., traveling speed, language spoken, ethnic group, and etc.) Also, in some embodiments, the column for the structure may not be necessary if all neural networks to be employed have the same structure.

In some preferred embodiments, the C-DBMS 409 can also include a process map for each setting. In various embodiments of the present invention, a process map can be a neural network workflow, a neural network schema, or a neural network descriptive document. In an example, a process map can include multiple ImNNs (each with a corresponding A) sets of coefficients and/or a neural network structure associated with the data values or B) the neural network executable modules (or pointers thereto) having the structure and/or the coefficients) connected serially, in parallel, or in combination with possible intermittent logic/algorithm, as illustrated with an example in Table 2 (that is, the “n” entry therein). In these preferred embodiments, the C-DBMS 409 query results in a process map. The SCDC 407 interprets the process map and instantiates neural networks in accordance with the process map.

TABLE 2

En-try #	Time	Loca-tion	Weath-er	Am-bient Temp	Structure	Coefficient Array
1	Day time range	1 st Ranges of Latitudes Longitudes	Sunny	Above freezing	feedforward	1 st set [. . .]
2	Day time range	2 nd Ranges of Lats + Longs	Cloudy	Above freezing	feedforward	2 nd set [. . .]
3	Night time	3 rd Ranges of Lats + Longs	Rain	Above freezing	Back propagation	3 rd set [. . .]
...						
n	Evening twilight	N th Ranges of Lats + Longs	Foggy	Above freezing	Map: Input to two feedforward neural networks Sum the output from the two nets then the sum to a back propagation	Array for the first feedforward Array for the second feedforward Array for the back propagation

In some embodiments, the entries that populate the C-DBMS 409 are made in such a way that there is i) no overlap between the possible setting characteristic values between different settings and ii) no null space between or outside the possible setting characteristic values between different settings. In these embodiments, when a query is made to the C-DBMS 409 by the SCDC 407 with the received setting characteristic values from the MTBC 403, one entry will be matched among the entries on the C-DBMS 409 and the information (e.g., A) the pointer to one matching neural network executable module or B) the values of a set of the coefficients and neural network structure thereof) will be sent back to the SCDC 407. An example of these embodiment is entries for 48 contiguous States—each entry defining the ranges of longitudes and latitudes for a State. In this embodiment output from a GPS device should fall into one of the 48 entries, and there is no null space between the ranges for the States. If the GPS is to operate within the 48 States, there is no null space outside thereof.

In some other embodiments, there can be some null spaces between or outside the possible setting characteristic values. An example of these embodiments is an entity training neural networks for automated machines that are to operate within large cities. Such a set of entries may have null spaces outside the large cities. In these embodiments a null value will be sent back to the SCDC 407, when a set of setting characteristic values falls into a null space. The SCDC 407 in turn can instruct the currently operational ImNN to continue to operate. There can be other instructions such as stop operating the entire automated machine, or send a signal for an augmented manual operation. More on the input sample data being outside the input sample space is described below in connection with FIG. 11.

In some other embodiments, there can be overlaps between the possible characteristic values. The overlaps can be partial or complete. If the C-DBM 409 is queried with a set of setting characteristic values that falls within such an

overlap, the C-DBMS 409 can return more than A) one neural network executable modules or B) one set of coefficients and neural network structures. An example of these embodiments is an entry for large cities and entry for the downtown of the large cities. These two entries could overlap. In these embodiments, the SCDC 407 can determine to use one of the more than one set returned from the C-DBMS 409. In one example, the SCDC 407 can use the set that cover the largest geographical area or use the set that cover the smallest geographical area. This feature of using geographical setting characteristic value ranges to address overlaps can be applied to other setting characteristics and/or a combination thereof.

In some embodiments, a neural network confidence level for each entry can be included as another column to, e.g., Table 1. The confidence level for each entry represents the confidence level for the neural network that is instantiated. For ease of reference, this confidence level is referred to as a neural network (NN) confidence level, which is different from an output confidence level. As noted above, an output confidence level is the confidence level of the selected class (i.e., output) for given input data being the correct one based on the scores of other classes. An NN confidence level can be determined based on probabilistic analysis of the training data set. For example, training data set having a narrow distribution among input sample values may be given a higher NN confidence level compared with another training data set having a broad distribution among its input sample values or vice versa depending on settings and/or applications. In another example, in some embodiments a training data set is associated with a testing data set. The NN confidence level can be the score of correct outcomes of a particular neural network after inferencing with such a testing data set.

In the embodiments with NN confidence levels, the entries returned by the C-DBMS 409 because of the overlap may also have the NN confidence levels. The SCDC 407 can use the values of the NN confidence levels, e.g., pick the entry with the highest NN confidence level.

It should be noted that different parts of the embodiments of the present invention can be implemented by different manufacturing entities. That is, the sensors and various components on the MC 405 can be manufactured by one or more entities, while the C-DBMS 409 entries can be populated by other manufacturing entities. In other words, this allows some manufacturers to concentrate on improving sensors and such, while allowing other manufacturers/entities to concentrate on improving the accuracies of ImNNs. For these example embodiments, the electronic format of the entries for the C-DBMS 409, the type of database management system used, and others may be specified (e.g., standardized) such that the C-DBMS 409 can be populated, queried, receive results of queries, and updated by different entities. Another aspect of these advantages of the present invention may be that the user of the SCDC 407 is allowed to test the accuracy of the entries in the C-DBMS 409 to accept or reject after testing. In some embodiments, an NN confidence level can be assigned to each of the entries in the C-DBMS 409.

Without storing the coefficients for different settings on a database management system, numerous neural networks can be deployed on the automated machine (i.e., the coefficients already fixed for each deployed neural network as in example embodiments described above in connection with Table 1). However, such arrangement requires numerous neural networks and may not be adaptable to new settings without updates to the deployed neural networks.

In some embodiments, the controllers—MC 405, SCDC 407, and C-DBMS 409 can be setup as standalone processes communicating with inter-process communication (IPC) protocols, as described in more detail below in connection with FIG. 5.

3. Plug-in Smart Architecture (PISA)

Various preferred embodiments described above can be implemented on a computing machine, for example, as a set of modules on a processor. FIG. 5 illustrates such an example preferred embodiment, which is described in terms of modules created using various memory spaces on a processor and in terms of various aspects of the operations of the modules. Each module can be considered as an individual machine when being operational (e.g., executing stored instructions) on a processor. In particular, a system module 501 (an example embodiment of the MTBC 403) may initiate a PISA module 503, which is an example implementation of the SCDC 409. This initiation can be performed by creating a fork 504 by a control process 573, which in some embodiments control the status and/or the operations of the PISA module 503 and the neural network module 507. The system module 501 can also initiate the input data stream 505 and initiate the output data stream 507 with a handle (e.g., a pointer) for the input stream 505 and another handle for the output stream 507, respectively.

The system module 501 created in memory space 503 can include interfaces to send/receive input/output to/from various sensors/devices such as a light-detecting and ranging radar (LIDAR) 551, global positioning system (GPS) 553, inertial measurement unit (IMU) 555, camera sensors 557 or the like. The system module 501 can have its own controlling algorithms relating to sensing 559 (receiving data from various input sensors/devices, perception 561 that analyzes the received data, decision 563 for making decisions based on the perceptions, and/or planning 565 to carry out the decisions. Output from all or subparts of the controlling algorithms can form a part of the input data stream 505, in addition to various sensors/devices with which the system module 501 is configured to interface. The various steps can be performed on real-time operating system (OS) 567 and on a Graphical Process Unit (GPU) 569 and/or Floating Points Graphical Accelerator (FPGA) 571.

The PISA module 503 may perform the following tasks:

- a. Gets the System State 513 (e.g., the values of the current setting characteristic values in FIG. 4) through, for example, the PISA Bus Library 508, which can be a library of interfaces that allows the PISA module 503 to interface with the Neural Network module 507 and the system module 501 in carrying out various functions/algorithms/routines as described in this Section.
- b. PISA Database Control 505 is configured to interface with (including querying) a Configuration Library 506 (e.g., C-DBMS 409). The Configuration Library 506 contains various neural network Coefficients Array and Structures (e.g., Table 1). Based on the System State values, the corresponding neural network coefficients and/or neural network structure(s) are retrieved from the Configuration Library 506. In turn, the retrieved neural network coefficients and/or neural network structure(s) are used to instantiate a neural network module 507.
- c. PISA Business Logic 509 can send/receive the status information of the neural network module 507 (e.g., the MC 405). The neural network module 507 can be created by a fork 508 in a new memory space as well as assigned with a listening socket 511. This socket 511

can be used to share current status of the neural network module 507 with the PISA module 503.

- d. PISA Bus Lib 508 may continually poll for the current System State

- i. If the System State changes to a different state (that is, the setting changes) from the previous state, the following may occur:

1. The PISA (Business Logic) 509, which contains various functions/algorithms/routines as described in this Section, initiates a kill process through the status socket or
2. The PISA (Business Logic) 509 initiates a suspend process through the status socket or
3. The PISA (Business Logic) 509 initiates an update system configuration process through the status socket.
4. Step b above may be followed by the above 1 or 2 or 3 processes.

- ii. If the System State does not change, no action is taken.

- e. PISA (Business Logic) 509 can review the status from the Neural Network module 507 and communicates status to the System module 501 via the PISA Bus Lib 508.

Neural Network module 507 (instantiated via a fork 510 from the PISA module 503) may perform the following tasks:

- a. The Neural Network module 507 initiates a neural network dynamically linked library (NN DLL) 517 (e.g., ImNN in FIG. 4) within the module's memory address space;
- b. The PISA Bus Lib 508 establishes connection to System Input Stream 505;
- c. The PISA Bus Lib 508 establishes connection to System Output Stream 507;
- d. PISA Xmitter 523 establishes the Status Socket 511 back to the PISA module 503;
- e. The NN DLL 517 may perform the following:
 - i. The NN DLL 517 processes data from System Input Stream 505 to generate results for System Output Stream 507 via the PISA Bus Lib 508 (that is, performs inferences on the Input Stream);
 - ii. NN DLL 517 pushes status via the PISA Xmitter 523 through the Status Socket 511 on regular intervals to the PISA module 503;
 - iii. NN DLL 517 processes requests from the PISA module 503 as requested to include:
 - 1) Kill current NN DLL 517 process
 - 2) Suspend current NN DLL 517 process
 - 3) Update system configuration

It should be noted that in some preferred example embodiments, the PISA Bus Lib 508, the PISA Bus Lib 521 and another PISA Bus Lib (not shown) on the system module 501 can be the same set of interface routines/managers. In other example preferred embodiments, the PISA Bus Lib on the system module can have the largest set, a subset of which is included in the PISA Bus Lib 508, and in turn a subset of which is included in the PISA Bus Lib 521.

FIG. 6 illustrates various preferred steps performed by the system module 501. In step 601, raw data (as received from the system module 501 but the system module 501 may have performed some operations the data) from various sensors/devices are collected. In step 603, system state is generated (predefined output data from various sensors and/or devices interfacing with the system module 501). In step 605, the stored system state becomes available to be shared with

other modules. In step 607, system input data is generated (predefined output data from various sensors and devices interfacing with the system module 501, and not necessarily the same data compared with the system state generated in step 603). In step 609, the input data stream is generated, which is the system input stream 505 in FIG. 5. In step 611, the output data stream (the system output stream 507 in FIG. 5) is obtained from the neural network module 507. In step 613, system output data is generated. In step 615, the generated system output data is processed to be used in controlling various components interfacing with the system module 501.

FIG. 7 illustrated various preferred steps performed by the neural network module 507, with step 701 as a starting point. In step 703, information is read for the neural network to be instantiated such as neural network coefficients and/or neural network structure(s). The neural network coefficients/structure(s) can be part of the signal for MC 417, illustrated in FIG. 4, and can be stored in memory as in step 705. A neural network can be initiated with the read the neural network coefficients/structure(s) in step 707—in the example of FIG. 5 only one neural network is instantiated that is designated as NN DLL and designated as ImNN in FIG. 4. Once initiated, the NN DLL starts its operations—including reading input data stream (step 709). The input data stream (step 753) is received from the system module 501, for example. The NN DLL then performs the function of inferencing (step 713) on the input data stream. After each inferencing performed on each set of input data stream, the NN DLL can issue a status to indicate the inference was normal or not.

When the status of the inferencing performed is checked in step 713, if it is not a normal operation (the branch marked with “-1” for step 713), the status is checked for an error in step 715. If there is error, the error code is written out, step 717. The error code is sent over the status socket 511 in step 755. If there is no error code, the status is checked for a warning code, step 719. If there is a warning code, the warning code is written out, step 721. The warning code is sent over to the status socket 511 in step 755. If there is no warning, the status is checked from information to be sent back to PISA 503. If there is information, the information is written out, step 725. The information is sent over to the status socket 511 in step 755.

When the status of the inferencing performed is checked in step 713, if it is a normal operation (the branch marked with “0” for step 713), an operational flag is checked (step 731). The operational flag set based on the “set action” 757 received from the PISA. If the operational flag is set to be on, the “Y” branch is taken and the NN wrapper writes the output of the instantiated neural network as output data stream (step 710) to be read by the system 501 (step 751). If the operational flag is set to be off, the “N” branch is taken and the NN is terminated in step 733, which ends the operation of the NN wrapper in step 735. Here, an example of an error code is generated when an unrecoverable error has occurred and the NN DLL 517 is to be terminated. An example of a warning code is generated when a recoverable error has occurred and a warning message is to be sent to the PISA module 503. An example of an information code is when the NN DLL 517 completes a task without an error.

FIG. 8 illustrated various preferred steps performed by the PISA module 503 with step 801 as a starting point. The box 800 depicts a set of system state values that are read when the Get System State 513 is performed. In step 803, the input system state (e.g., setting characteristic values) is read. The input system state includes the set of system state values

(box 800) and status information from the NN wrapper in step 755. If it is determined that a new set of neural network coefficients and/or structure are needed based on the input system state, then in step 805 is performed—that is, the values of the input system state is sent (in step 807) to the coefficient DBMS (step 809) which outputs the neural network coefficients and/or neural network structure associated with the input system state, in step 811. An instruction to instantiate is sent over to the neural network module (in step 813) along with the aforementioned coefficients and/or structure information 815. The neural network module then sends a status information 817, which is received in step 819.

When the status is checked in step 821, if it is not a normal operation (the branch marked with “-1” for step 821), the status is checked for an error in step 823. If there is error, the error code is written out, step 825. The error code is written as a system status in step 827. If there is no error code, the status is checked for a warning code, step 823. If there is a warning code, the warning code is written as a system status in step 827. If there is no warning, the status is checked from information code. If there is information code, the information is written as a system status in step 827. Here, an example of an error code is generated when an unrecoverable error has occurred and the NN DLL 517 is to be terminated. An example of a warning code is generated when a recoverable error has occurred and a warning message is to be sent to the PISA module 503. An example of an information code is when the NN DLL 517 completes a task without an error.

The system status is interpreted to determine an action in step 829, and the determined action is sent in 831 to determine if the neural network module is to continue to inference—and to the NN Module. If it is determined to continue, the PISA module continues to execute. If it is determined to terminate, then the PISA module is terminated.

The pseudo-computer program provided in the section below is an example preferred implementation of the present invention. In particular, PISAController performs the following steps:

a) determines, using the command line arguments, the NN implementation class name and potentially a setting/coefficient set database implementation class name.

b) initializes a class called PISANN. Note that the implementation class name which is capable of both inheriting the PISANN class and implementing the PISAInterface. This ensures that access to the methods is available to the PISAController main section and ensures that the correct methods are completely implemented by the dependent section(s).

c) dynamically loads the primary NN and if specified, the setting/co-efficient set database. The resulting dynamic allocation does not require pre-compiled knowledge of the class. Also, PISAController has no need to have the insight into the inter-workings of the dynamically loaded neural network (NN).

d) determines if the PISANN has been pre-trained or trained on instantiation. If not, the main class is exited since the incoming NN needs to be trained prior to execution in this embodiment.

e) determines if the PISANNDatabase is available and connects to the NN database. If not available, in this embodiment, a default database is provided to the PISANN-Database for use during operations.

f) initiate PISAInputDataHandler based on the command line arguments. All input data can be pulled directly from the

PISAInputDataHandler. This class can be modified to support multiple sources (e.g. database, files, real-time feeds, sensor feeds).

g) initiate PISAOutputDataHandler based on the command line arguments. All output data can be pushed directly into the PISAOutputDataHandler. This class can be modified to support multiple sources (e.g. database, files, real-time feeds, sensor feeds).

h) initiate PISAEvent based on the command line arguments. This class can be modified to support multiple system event types (e.g. change in settings from location-based sensors or GPS, change in settings from temperature sensors, etc.). The PISAEvent can be evaluated at any time during system processing. In this embodiment, the PISAEvent is checked on each iteration of new input data.

i) runs through input data gathered from the PISAInputDataHandler performing inference with the PISANN where inference results are placed into the PISAOutputDataHandler. For each iteration, the PISAEvent is checked for an updated status. If a PISAEvent has a changed status, new coefficients are retrieved from the PISANNDatabase, applied to the PISANN, and processing continues with the current PISAInputDataHandler and PISAOutputDataHandler. If no coefficients are available, the PISAController terminates. In this example embodiment, the PISAController continues to run until the data provided stops being produced from the PISAInputDataHandler.

```

/* Main driver program for TOPController */
import java.lang.reflect.Constructor;
import java.lang.reflect.Method;
import java.util.*;
public class PISAController {
public static void main(String[] args) {
// Initialize the input NN, NN database
String inputNNClassName = null;
String inputNNDatabaseName = null;
boolean hasDatabase = false;
// Check the main arguments
if (args.length > 2) {
inputNNClassName = new String(args[1];
inputNNDatabaseName = new String(args[2]);
hasDatabase = true;
} else if (args.length > 1) {
inputNNClassName = new String(args[1];
inputNNDatabaseName = new String("DDB.class");
} else {
System.out.println("PISA requires class name.");
System.exit(-1);
}
try {
// Initialize primary PISA NNs
PISANN pisaNN = null;
PISANNDatabase databaseNN = null;
// Generate the PISA NN object from the ClassLoader
PISAClassLoader pisaCL = new PISAClassLoader ();
pisaNN = pisaCL.invoke(inputNNClassName,
"getNN");
// Get default trained NN
if (!(pisaNN.getNNtrained( ))) {
System.out.println("PISA requires a trained network.");
System.exit(-1);
}
// If PISAController configuration has a database,
// create an instance of database and connect.
// Generate PISANNDatabase object from the ClassLoader
PISAClassLoader pisaCL = new PISAClassLoader ();
databaseNN = pisaCL.invoke(inputNNDatabaseName,
"getDB");
// Connect to the database using the default class name
// which contains the name of the database connection.
try {
databaseNN.connectDB();
}
}
}

```

-continued

```

catch (Exception e) {
e.printStackTrace();
System.exit(-1);
}
// Initialize PISA input data
PISAInputDataHandler inputData = new
PISAInputDataHandler(args[0]);
// Check on the input data class
if (inputData == null) {
System.out.println("PISA requires an input data
stream.");
System.exit(-1);
}
// Check on size of input data stream.
if (inputData.getInputSize( ) <= 0) {
System.out.println("PISA requires an input data
stream.");
System.exit(-1);
}
double[] inputs = new double[inputData.getInputSize( )];
// Initialize PISA output data
PISAOutputDataHandler ouputData = new
PISAOutputDataHandler(args[0]);
// Check on the output data class
if (ouputData == null) {
System.out.println("PISA requires an output data
stream.");
System.exit(-1);
}
// Check on size of input data stream.
if (outputData.getOutputSize( ) <= 0) {
System.out.println("PISA requires an output data
stream.");
System.exit(-1);
}
double[] outputs = new double[ouputData.getOuputSize( )];
int cycleIndex = 0;
// Initialize PISA settings change events
PISAEvent pisaEvent = new
PISAEvent (args[0]);
// Initialize output data
double[] outputNN = null;
// Continually loop over input data
// Output results to the outputData handler.
// until end of input data or
// until pisaEvent has changed
// if pisaEvent, then
// get new coefficients
// if new coefficients possible
// update the NN
// else
// terminate NN and exit
// continue
while (cycleIndex >= 0) {
// Get input data
inputs = inputData.getInput(cycleIndex);
try {
// Perform inference using pisaNN
outputs = pisaNN.inferenceNN(inputs);
// Perform setting check
if (pisaEvent.hasChanged( )) {
// Pull new coefficients based on the new event
// apply the coefficients to the NN
double[] coefficients;
try {
coefficients = databaseNN.getCoefficients(
pisaEvent.getEvent( ),
pisaEvent.getEventDetails( ));
}
}
catch (Exception e) {
// No coefficients found for event and
// settings change
e.printStackTrace();
System.exit(-1);
}
}
try {
pisaNN.setCoefficients(coefficients);
}
}
catch (Exception e) {
}
}

```

-continued

```

// Coefficients are not a match for NN
// Or error applying coefficients to the
NN
    e.printStackTrace( );
    System.exit(-1);
    }
    } else {
        outputData.setOutput(outputNN);
    }
    }
    catch (Exception e) {
        e.printStackTrace( );
    }
    cycleIndex++;
    if (cycleIndex > inputData.getCycleMax( )) {
        cycleIndex = 0;
    }
    }
    }
    }
    // Catch all errors.
    catch (Exception e) {
        e.printStackTrace( );
        System.exit(-1);
    }
}

```

4. Boundary Conditions of Input and Output Spaces

FIG. 9a graphically illustrates a simplified decision-making space 301 that shows both the input data space and output results from neural networks. In particular, outer polygonal boundary 303 may depict the entire input sample space (e.g., the decision-making space) in two dimensions, and two smaller circles, 305 and 307, located therein may depict validated output classes. A neural network can be constructed (e.g., instantiated having a given set of coefficients and a particular neural network structure) and trained using sample input data, either supervised or unsupervised, to classify input data into output categories. It should be noted that in some preferred embodiments, output can be generated from a node(s) of an output layer or a node(s) from a layer between an input layer and an output layer. Here, structuring a neural network includes selecting an appropriate neural network structure (e.g., convolutional neural network, CNN) and providing an adequate number of nodes and layers of nodes for the given classification goal. In one example embodiment, the sample space may represent two features from a set of images, the polygon 303 representing the entire range in which values of those two features can take for the sample images. Continuing with the example, one class 305 can be images that have pictures of cats, and the other class 307 can be images that have pictures of dogs. Here, a neural network can be constructed and trained using training sample images to inference (e.g., classify) whether an image contains a picture of a cat, a dog, both, or neither. The input space bound by 303 can be considered for the entire sample space of one particular setting. In one example, a different setting may have input sample space that does not overlap with the space 303. Under this example, the input samples themselves may indicate a different setting that requires A) a new neural network executable module or B) a new set of coefficients and/or neural network structure—that is, when examining input samples, if it is outside the sample space for a given setting, this may indicate a need to query a database (e.g., C-DBMS 409). In one example, the input space 303 may represent the entire input space for images collected in a desert setting. In such an example, when the setting changes

to a forested region, the images collected in the forested region may not fall within the input space 303.

In some embodiments of the present invention, boundary conditions in the output space is used in operating/control-
 5 ling neural networks, ImNNs. In connection with FIG. 9b, a set of boundary conditions can be described as allowing the output of a neural network to be only within a certain range—e.g., Region A 351, although the input data can be anywhere within the entire sample space as depicted in FIG. 9a. Referring back to FIG. 9b, if an output from a neural
 10 network constructed and trained to inference classes located within Region A, the output can be used in a subsequent processing, described below in connection with FIG. 12 below. However, if output of such a neural network is
 15 outside of Region A (e.g., Region B 353 or Region C 355 of FIG. 9b), the output can be discarded and not used. In another simplified depiction of FIG. 10a, the decision-making can be illustrated as a function in a one-dimensional space. In this simplified version, the boundary conditions are
 20 depicted as a range 1001 in which an output from a neural network is checked against.

Continuing on with the above output space description, in a simplified example, a neural network 1003 structured to
 inference input data 1002 to generate output can be instan-
 25 tiated. The output can be checked to determine against the output breach boundary cognition(s). If “no,” the output is forwarded to the next step 1007 to be used by a machine to be controlled (e.g., MTBC 403). If “yes,” this can be considered an event to query the C-DBMS and/or the output
 30 is not forwarded to the next step.

The step of determining the severity of breaching the boundary conditions can be illustrated in connection with
 FIG. 9b. That is, in some embodiments of the present invention, multiple sets of boundary conditions can be
 35 imposed. In the preferred example embodiment of shown in FIG. 9b, three regions are shown. The first region is referred to as Region A 351, in which the output from a neural network would have been forwarded to the next step in the processing chain. Output falling within Region A 351 is
 40 considered as not breaching the boundary conditions and/or satisfying the boundary conditions. The second region is referred to as Region B 353, in which the output could be considered as breaching (e.g., violating or exceeding) the boundary conditions but not harmful to the machine or
 45 anyone/anything surrounding the machine. In this case, the output can be ignored/discarded and not forwarded to the next step in the processing step. The third region is referred to as Region C 355, in which the output breaches the boundary conditions to such an extent that it could cause
 50 harm to the machine or to someone/something surrounding the machine. In such a case, the machine can be shut down immediately or the user can be notified that the machine needs to be used in its manual mode. In another example embodiment of the present invention, if the boundary condition is severely breached as Region C 355, a presumption
 55 can be made that the machine to be controlled is in a new setting. In this example, a database (e.g., the C-DBMS 409) can be queried to obtain A) a set of coefficients and/or a neural network structure associated with the received data
 60 values or B) the neural network executable module (or a pointer thereto) having the structure and/or the coefficients from the database that may match with the current setting characteristic values.

In an exemplary embodiment, a speech generator can be
 65 equipped with various features of the present invention. In particular, an exemplary preferred speech generator can be coupled to a user identifier such as a speech recognition

system. Initially, the speech generator can be set to generate using a default setting (e.g., the predominant language of the geographical location in which the generator is placed) or a previous setting (e.g., the language spoken by a previous user). During the operation, the speech recognition system can be configured to determine the speech of the current user. If the language used by the current user is different from the default/previous setting (that is, outside the input sample space for the predominant language or the language of the previous user), the speech recognition system can be further configured to identify the language the user (e.g., English, German, French, etc.). If the user is speaking in a language different from the default/current setting, the C-DBMS 409 can be queried for the user's language, selected and loaded for generating speech in the language of the user. In some embodiments, the C-DBMS 409 can be queried for each new user.

Similarly, a facial recognition system can be set to identify a user by using a default setting (e.g., the predominant ethnic group in the geographical location in which the facial recognition system is placed) or a previous setting (e.g., the ethnic group of a previous user). During the operation, the facial recognition system can be configured to determine the ethnic background of the current user. If the ethnic background of the current user is different from the default/previous setting (that is, outside the input sample space for the predominant ethnic group or the ethnic group of the previous user), the facial recognition system can be further configured to identify the ethnic background of the current user. If the current user belongs to an ethnic group different from the default/current setting, the C-DBMS 409 can be queried for the current user's ethnic group, selected and loaded for facial recognition. In some embodiments, the C-DBMS 409 can be queried for each new user. A neural network trained with training data set for a narrowly defined setting (e.g., ethnic groups for facial, language, and/or regional accents in speaking languages) may yield more accurate results than a neural network trained with broad, disparate settings.

Some embodiments of the speech generator may include an implementation neural network constructed and trained to generate signals/data that can become human understandable phrases, sentences, and etc. when played on a loudspeaker. That is, when the ImNN of the speech generator outputs one of forbidden words, the trigger event detector recognizes it as a forbidden word (e.g., outside output boundary condition), and does not forward the output of the speech generator to a loudspeaker and/or terminates the currently running ImNN and instantiates a new ImNN having a different set of coefficients and/or different neural network structure.

Although boundary conditions have been illustrated in connection with one-dimensional decision space, two-dimensional decision space, speech generation, facial recognition contexts, the use of boundary conditions can be also expressed in terms of triggering events (that is a triggering event being a form of breaching a boundary condition), in terms of hard operating limitations of the machine being controlled, and/or in terms of using output confidence levels of the outputs of neural networks for given settings. In addition to expressing boundary conditions as triggering events, boundary conditions can also be viewed as expressions of the competence range in which a given neural network is constructed and trained to operate per a particular setting. Also, a different way to define boundary conditions can be in term of the output confidence level in connection with a given output from a neural network. In one example

preferred embodiments, if the output confidence level of an output of a neural network falls below a predetermined level (e.g., below 60%), such an output can be discarded and/or A) a new neural network executable module or B) a new set of coefficients and/or structure can be searched and selected. In another example preferred embodiments, if the output confidence levels of two or more outputs of a neural network are similar (e.g., the same or only different marginally as in less than 5%), such a set of outputs can be discarded and/or A) a new neural network executable module or B) a new set of coefficients and/or structure can be searched and selected.

5. Triggering Event Detector

As shown in FIG. 11, in some preferred embodiments of the present invention, a triggering event detector (TED) 1131 is included. In these embodiments, an MTBC 1103, an SCDC 1107, an MC 1105, an ImNN 1121, and a C-DMBS 1109 have features/functions/capabilities of the MTBC 403, the SCDC 407, the MC 405, the ImNN 421, and the C-DMBS 409, respectively, as described above in connection with FIG. 4. Also coupling mechanisms can be included such as: input data 1119 (two paths shown in FIG. 11), control data 1115 (two shown in FIG. 11), status data 1113, a signal for MC 1117 have features/functions/capabilities input data 419, control data 415, status data 413, a signal for MC 417, respectively, as described above in connection with FIG. 4. In addition, the MTBC 1103, the SCDC 1107, the MC 1105, the ImNN 1121, and the C-DMBS 1109 (and the coupling mechanism) are configured to work with the TED 1131 as described below.

The TED 1131 receives the input data from the MTBC 1103 and control data from the MC 1105. In various embodiments of the present invention, the input data and control data sent to TED 1131 can be synchronized. That is, the input data to the MC 1105 that caused certain control data to be generated by the MC 1105 after a process delay can be sent to the TED 1131 at the same time (or associated with each other) to be processed by the TED 1131. A triggering event can relate to input sample(s) being detected to be outside the input sample space for a particular setting and/or output data breaching the boundary conditions (either for a particular setting or a universal breach). In FIG. 11, if a triggering event is detected, a signal 1133 send a notice to the MTBC 1101 and/or SCDC 1107. The notice can be an instruction to discard the output (synchronized with the input that caused the triggering event), or the notice can be an instruction to lower the NN confidence level that is instantiated on the ImNN 1121 (e.g., the SCDC 1107 notifies the CDMS 1109 to store the lowered confidence level for the corresponding A) neural network executable module or B) set of coefficients and/or its structure. In another preferred embodiment, upon receiving such a signal, the MTBC 1103 collects the setting characteristic values and sends them to the SCDC 1107, which in turn queries the C-DBMS 1109. In another preferred embodiment, upon receiving such a signal from the TED 1131, the SCDC 1107 uses the setting characteristic values to query the C-DBMS 1109.

As such, a trigger event detector is an example of mechanism(s) in detecting/sensing boundary conditions. In some embodiments, the triggering event detector is implemented using a neural network that is constructed and trained to detect one or more of triggering events or a type of events. In other embodiments of the present invention, a set of logical steps in algorithms/heuristics can be used to detect one or more triggering events or a type of events. In some preferred embodiments, similar to the input sample space, the output space can also be defined by range of values. In these embodiments, logic to detecting a triggering event

determines if the control data (i.e., output of the MC 1105) is outside the predefined output space. In yet some embodiments, the TED 1131 can have a neural network and a set of logical steps.

TABLE 3

Entry #	Time	Location	Weather	Ambient Temp	Input Sample space	Boundary Conditions	Structure	Coefficient Array
1	Day time range	1 st Ranges of Latitudes Longitudes	Sunny	Above freezing	Input ranges	Output ranges AND/OR structure & coefficient array	feedforward	1 st set [. . .]
2	Day time range	2 nd Ranges of Lats + Longs	Cloudy	Above freezing	Input ranges	Output ranges AND/OR structure & coefficient array	feedforward	2 nd set [. . .]
3	Night time	3 rd Ranges of Lats + Longs	Rain	Above freezing	Input ranges	Output ranges AND/OR structure & coefficient array	Back propagation	3 rd set [. . .]
...								
n	Evening twilight	N th Ranges of Lats + Longs	Sunny	Freezing	Input ranges	Output ranges AND/OR structure & coefficient array	Restricted Boltzmann	N th set [. . .]

As illustrated in Table 3 above, in some preferred embodiments, setting characteristic value ranges can also be associated with input sample space (e.g., defined by the ranges of input sample values) and output boundary conditions. In particular, an input space ranges is defined for a given neural network of Coefficients Array and Structure. The input ranges are used as described in connection with FIG. 11 (and not necessarily the output boundary conditions in that figure). Various preferred embodiments described below in connection with FIG. 12, the input boundary conditions, and output boundary conditions may be used.

In particular, the preferred embodiments in connection with FIG. 12 may include an MTBC 1203, an SCDC 1207, an MC 1205, an ImNN 1221, and a C-DMBS 1209 have features/functions/capabilities of the MTBC 1103, the SCDC 1107, the MC 1105, the ImNN 1121, and the C-DMBS 1109, respectively, as described above in connection with FIG. 12. Also coupling mechanisms can be included such as: input data 1219 (two paths shown in FIG. 12), control data 1215 (two shown in FIG. 12), status data 1213, a signal for MC 1217 have features/functions/capabilities input data 419, control data 415, status data 413, a signal for MC 417, respectively, as described above in connection with FIG. 4. In addition, the MTBC 1103, the SCDC 1107, the MC 1105, the ImNN 1121, and the C-DMBS 1109 (and the coupling mechanism) are configured to work with the TED 1131 as described below. In addition, the MTBC 1203, the SCDC 1207, the MC 1205, the ImNN 1221, and the C-DMBS 1209 (and the coupling mechanism) are configured to work with the TED Data 1235 and data on TED Output 1233 as described below.

While A) the neural network executable module or B) the coefficients and structure are used by the SCDC 1207 to instantiate the ImNN 1221 in the MC 1205 for a particular set of setting characteristic value ranges, the corresponding input space ranges and output boundary conditions can be loaded on to the TED 1231 (via the TED Data lines 1235

shown in FIG. 12). In the embodiments that use a neural network, the coefficient set and the neural network structure to be used in the TED 1231 are also sent to the TED 1231. Upon receiving the data from the SCDC 1207, TED 1231

can implement specific configurations. Preferred exemplary embodiments contemplate, triggering event as:

Incorrect/abnormal type: Output(s) being out of operating bounds/limitations—examples:

In a refrigerator controller, the controller attempts to raise the temperature of fridge above the recommended operational temperature

In an oven the controller attempts to raise the temperature above a recommended operational temperature

In an oven and/or stove, turn on the oven and/or stove during a time or condition when it has been designated for non-use (for example, between 1 am and 6 am, or when no one is at home, e.g., when a sensor determines that no one is home)

In a speech generator, curse words or other inappropriate words are generated

In a controller for a driverless car, the controller issues a lane change command after receiving a proximity warning

In an image generator, inappropriate images are generated
In an image display apparatus, inappropriate images are displayed

In a controller for a robot, a command to harm a human being is created

In a printer controller, a counterfeit currency or counterfeit signature is generated

Security Breach Type:

In a controller for a refrigerator (or another system for example a video camera system) repeated information requests to particular websites

An authorized access to personal information

An attempt to adjust or replace or otherwise modify the controller

An attempt to cause a denial of service attack on a remote device

While running two virtual machines on an autonomous machine with substantially identical ImNNs on each

virtual machine, one set of ImNNs start generating output data deviating from the output of the other set of ImNNs

Unauthorized usage level type: In an automated personal assistant embodiment, when a user is assigned to a G-rated search results only, the personal assistant generates results that are in R-rated category.

Going back to FIG. 12, when a triggering event is detected, the TED 1231 sends a signal over TED Output 1233 to the MTBC 1203 and/or the SCDC 1207. This signal indicates that a triggering event is detected and an action is required to be taken. In one preferred embodiment, upon receiving such a signal, the MTBC 1203 collects the setting characteristic values and sends them to the SCDC 1207, which in turn queries the C-DBMS 1209. In another preferred embodiment, upon receiving such a signal from the TED 1231, the SCDC 1207 uses the setting characteristic values to query the C-DBMS 1209. In other embodiments, if a triggering event is detected, TED Output 1233 send a notice to the MTBC 1201 and/or SCDC 1207. The notice can be an instruction to discard the output (synchronized with the input that caused the triggering event), or the notice can be an instruction to lower the NN confidence level that is instantiated on the ImNN 1221 (e.g., the SCDC 1207 notifies the CDMS 1209 to store the lowered confidence level for the corresponding A) neural network executable module OR B) the coefficients and/or its structure.

In some preferred embodiments, when such a triggering event signal is received the SCDC 1207 keeps the information about the entry of the C-DBMS 1209 that caused the triggering event. SCDC 1207 then updates that entry in the C-DBMS 1209. The updates can include lowering the NN confidence level of the entry (if the entry has an NN confidence level column as described above in connection with the C-DBMS 1209), remove the entry, and/or mark it for evaluation manually off-line.

Various components/devices of the SCDC and MC (described above in connection with FIGS. 4, 11 and 12) can be implemented on a chip, a chip-set, ASIC, AI server (e.g., DGX-1 by Nvidia), and/or firmware. This is, for example, to prevent a potential security breach (e.g., a virus attack) and/or to provide a baseline from which to re-boot. In other words, in some exemplary embodiments, the logic and/or neural network(s) located in the SCDC is not modifiable or adjustable by or at the autonomous machine, but only re-deployable, modifiable, and/or adjustable by an authorized system of the original manufacturer of the autonomous machine. It should be noted in some embodiments, such a SCDC can run on one thread (e.g., on one virtual machine), while ImNN(s) can run on another thread (e.g., another virtual machine) on a general-purpose processor or a graphical accelerator/processor (e.g., implemented on solid-state devices such as a chip, a chip-set, ASIC).

In various embodiments of the present invention, the SCDC and the ImNN(s) can be co-located on a device (e.g., a general-purpose computer, a controller chassis, an ASIC, chipset, etc.). Although the implementation of some of the preferred embodiments are described in terms of solid-state devices (e.g., semiconductor chips), portions of some preferred embodiments being implemented on an optical computer device or quantum computing device is also contemplated. It should be noted that the SCDC can also be implemented on an AI server (for example, DGX-1 by Nvidia), and/or firmware deployed on a server computer, a processor specifically adapted to allow efficient running of neural networks also referred to as neural network processors. The ImNN(s) can also run on a processor (e.g., a

general-purpose processor, or graphical accelerator/processor, digital processor or processors specifically adapted to allow efficient running of neural networks also referred to as neural network processors). As noted above, the SCDC can be implemented (e.g., on a server) remotely located from the ImNN(s) (e.g., on a client(s)).

In some embodiments of the present invention, the structure(s) of the ImNN(s) are not modifiable once deployed on an automated machine, for security reasons and/or for efficiency. In such an embodiment, only the coefficients for the nodes are stored in the C-DMBS and would be used by the SCDC to modify the ImNN(s). In other words, the information relating to the structures (e.g., type of neural network, number of nodes and layers, and nodal connection information) is not needed to be stored in the C-DMBS for these embodiments, since the neural network structure of the ImNN(s) is not modifiable. The structures of the ImNN(s) for these embodiments can be implemented on fixed hardware/firmware that cannot be changed once deployed.

Any module, routine or any apparatus configured to perform the functions recited by means described herein or may be performed by any suitable means capable of performing the corresponding functions. The means may include various hardware and/or software component(s) and/or module(s), including, but not limited to, a circuit, an application specific integrated circuit (ASIC), or processor. Further, it should be appreciated that modules and/or other appropriate means for performing the methods and techniques described herein can be downloaded and/or otherwise obtained by a user terminal and/or base station as applicable. For example, such a device can be coupled to a server to facilitate the transfer of means for performing the methods described herein. Alternatively, various methods described herein can be provided via storage means (e.g., RAM, ROM, a physical storage medium such as a compact disc (CD) or floppy disk, etc.), such that a user terminal and/or base station can obtain the various methods upon coupling or providing the storage means to the device. Moreover, any other suitable technique for providing the methods and techniques described herein to a device can be utilized.

As used herein, the term “determining” encompasses a wide variety of actions. For example, “determining” may include calculating, computing, processing, deriving, investigating, looking up (e.g., looking up in a table, a database or another data structure), ascertaining and the like. Further, “determining” may include receiving (e.g., receiving information), accessing (e.g., accessing data in a memory) and the like. In addition, “determining” may include resolving, selecting, choosing, establishing and the like.

Also, as used herein, phrases neural network executable modules, executable modules of neural network, executable neural network modules mean the same.

The various illustrative logical blocks, modules, processors and circuits described in connection with this disclosure may be implemented or performed with a general purpose processor, a digital signal processor (DSP), an application specific integrated circuit (ASIC), a field programmable gate array signal (FPGA) or other programmable logic device (PLD), discrete gate or transistor logic, discrete hardware components or any combination thereof designed to perform the functions described herein. A general-purpose processor may be a microprocessor, but in the alternative, the processor may be any commercially available processor, controller, microcontroller or state machine. A processor may also be implemented as a combination of computing devices, e.g., a combination of a DSP and a microprocessor, a plurality of

microprocessors, one or more microprocessors in conjunction with a DSP core, or any other such configuration.

As one of skill in the art will appreciate, the steps of a method or algorithm described in connection with the present disclosure may be embodied directly in hardware, in a software module executed by a processor, or in a combination of the two. A software module may reside in any form of storage medium that is known in the art, including memory that may be part of a microprocessor or in communication with a microprocessor. Some examples of storage media that may be used include, but are not limited to, random access memory (RAM), read only memory (ROM), flash memory, erasable programmable read-only memory (EPROM), electrically erasable programmable read-only memory (EEPROM), registers, a hard disk, a removable disk including removable optical media, and so forth. A software module may comprise a single instruction, or many instructions, and may be distributed over several different code segments, among different programs, and across multiple storage media. A storage medium may be coupled to a processor such that the processor can read information from, and write information to, the storage medium. In the alternative, the storage medium may be integral to the processor.

The methods disclosed herein may include one or more steps or actions for achieving a described method. The method steps and/or actions may be interchanged with one another without departing from the scope of the invention. In other words, unless a specific order of steps or actions is specified, the order and/or use of specific steps and/or actions may be modified without departing from the scope of the disclosure. The functions described may be implemented in hardware, software, firmware, or any combination thereof. If implemented in hardware, an example hardware configuration may comprise a processing system in a device. The processing system may be implemented with a bus architecture. The bus may include any number of interconnecting buses and bridges depending on the specific application of the processing system and the overall design constraints. The bus may link together various circuits including a processor, machine-readable media, and a bus interface. The bus interface may be used to connect a network adapter, among other things, to the processing system via the bus. The network adapter may be used to implement signal processing functions. For certain aspects, a user interface (e.g., keypad, display, mouse, joystick, etc.) may also be connected to the bus. The bus may also link various other circuits such as timing sources, peripherals, voltage regulators, power management circuits, and the like, which are well known in the art, and therefore, will not be described any further.

The processor (e.g., image processor) may be responsible for managing the bus and general processing, including the execution of software stored on the machine-readable media. The processor may be implemented with one or more general-purpose and/or special-purpose processors. Examples include microprocessors, microcontrollers, DSP processors, and other circuitry that can execute software. Software shall be construed broadly to mean instructions, data, or any combination thereof, whether referred to as software, firmware, middleware, microcode, hardware description language, or otherwise. Machine-readable media may include, by way of example, random access memory (RAM), flash memory, read only memory (ROM), programmable read-only memory (PROM), erasable programmable read-only memory (EPROM), electrically erasable programmable read-only memory (EEPROM), registers, magnetic disks, optical disks, hard drives, or any other suitable storage

medium, or any combination thereof. The machine-readable media may be embodied in a computer-program product. The computer-program product may comprise packaging materials.

In a hardware implementation, the machine-readable media may be part of the processing system separate from the processor. However, as those skilled in the art will readily appreciate, the machine-readable media, or any portion thereof, may be external to the processing system. By way of example, the machine-readable media may include a transmission line, a carrier wave modulated by data, and/or a computer product separate from the device, all which may be accessed by the processor through the bus interface. Alternatively, or in addition, the machine-readable media, or any portion thereof, may be integrated into the processor, such as the case may be with cache and/or general register files. Although the various components discussed may be described as having a specific location, such as a local component, they may also be configured in various ways, such as certain components being configured as part of a distributed computing system.

In some embodiments, the processing system may be configured as a general-purpose processing system with one or more microprocessors providing the processor functionality and external memory providing at least a portion of the machine-readable media, all linked together with other supporting circuitry through an external bus architecture. In some embodiments, the processing system may be implemented with an application specific integrated circuit (ASIC) with the processor, the bus interface, the user interface, supporting circuitry, and at least a portion of the machine-readable media integrated into a single chip, or with one or more field programmable gate arrays (FPGAs), programmable logic devices (PLDs), controllers, state machines, gated logic, discrete hardware components, or any other suitable circuitry, or any combination of circuits that can perform the various functionality described throughout this disclosure. In some embodiments, the processing system may comprise one or more neuromorphic processors for implementing the neuron models and models of neural systems described herein. As another alternative, the processing system may be implemented with an application specific integrated circuit (ASIC) with the processor, the bus interface, the user interface, supporting circuitry, and at least a portion of the machine-readable media integrated into a single chip, or with one or more field programmable gate arrays (FPGAs), programmable logic devices (PLDs), controllers, state machines, gated logic, discrete hardware components, or any other suitable circuitry, or any combination of circuits that can perform the various functionality described throughout this disclosure. Those skilled in the art will recognize how best to implement the described functionality for the processing system depending on the particular application and the overall design constraints imposed on the overall system.

The machine-readable media may comprise a number of software modules. The software modules include instructions that, when executed by the processor, cause the processing system to perform various functions. The software modules may include a transmission module and a receiving module. Each software module may reside in a single storage device or be distributed across multiple storage devices. By way of example, a software module may be loaded into RAM from another storage medium when a triggering event occurs. During execution of the software module, the processor may load some of the instructions into cache to increase access speed. When referring to the

functionality of a software module below, it will be understood that such functionality is implemented by the processor when executing instructions from that software module.

Some embodiments may comprise a computer program product for performing the operations presented herein. For example, such a computer program product may comprise a computer-readable medium having instructions stored (and/or encoded) thereon, the instructions being executable by one or more processors to perform the operations described herein. If implemented in software, functions may be stored or transmitted over as one or more instructions or code on a computer-readable medium. Computer-readable media include both computer storage media and communication media including any medium that facilitates transfer of a computer program from one place to another. A storage medium may be any available medium that can be accessed by a computer. Thus, in some embodiments a computer-readable media may comprise non-transitory computer-readable media (e.g., tangible media). Combinations of the above should also be included within the scope of computer-readable media.

What is claimed is:

1. A method of controlling a machine, the method comprising:

storing at least two sets of neural network coefficients, each being different from the others;

associating each of the at least two sets of neural network coefficients with one or more characteristics of a setting;

receiving first data from one or more input devices of the machine;

selecting one from the at least two sets of neural network coefficients based on the first data and the one or more characteristics of settings;

instantiating a neural network with the selected one from the at least two sets of neural network coefficients; and controlling an aspect of the machine using an output from the instantiated neural network.

2. The method of claim 1, wherein said each of the one or more characteristics of a setting is defined with a range of values.

3. The method of claim 2, further comprising storing information relating to a neural network structure associated each of the at least two sets of neural network coefficients.

4. The method of claim 3, wherein the neural network structure is one of a convolutional neural network, a feed forward neural network, a neural Turing machine, Hopfield neural network, or a Boltzmann machine neural network.

5. The method of claim 1, wherein the setting is one of a temperate urban region, a desert rural region, a forested mountain region, and a coastal city.

6. The method of claim 1, wherein selecting one from the at least two sets of neural network coefficients further comprises matching the first data with the one or more characteristics of settings.

7. The method of claim 6, wherein said matching further comprises:

comparing the first data with the one or more characteristics of settings, wherein each of the one or more characteristics of settings is defined with a range of values; and

identifying the selected one of the one or more characteristics of settings that the first data fall within the range of values.

8. The method of claim 6, wherein the neural network coefficients matched with the selected one are generated by using training data set collected within the corresponding particular setting.

9. The method of claim 1, wherein the first data includes data from a Global Positioning System.

10. The method of claim 1, wherein the information relating to the at least two sets of neural network coefficients is stored in a standardized format to allow access by electronic devices manufactured by different manufacturers.

11. The method of claim 1, further comprising:

storing a set of one or more input range values associated each of the at least two sets of neural network coefficients;

comparing the first data with the one or more input range values associated with the selected one from the at least two sets of neural network coefficients; and

selecting a new set among the at least two sets of neural network coefficients if the first data is outside the input range values.

12. The method of claim 1, further comprising:

storing a set of one or more output range values associated each of the at least two sets of neural network coefficients;

comparing the output with the one or more output range values associated with the selected one from the at least two sets of neural network coefficients; and

selecting a new set among the at least two sets of neural network coefficients if the output is outside the output range values.

13. An apparatus for controlling a machine, comprising: a database management system storing at least two sets of neural network coefficients being different from each other, at least one setting having one or more characteristics, and each of the at least two sets of neural network coefficients being associated with the at least one setting having one or more characteristics; and a controlling device that is coupled to receive first data from one or more input devices of the machine, arranged to select one from the at least two sets of neural network coefficients based on the first data and at one least one setting having one or more characteristics, and arranged to instantiate a neural network with the selected one from the at least two sets of neural network coefficients, wherein the neural network is configured to generate an output being used to control an aspect of the machine.

14. The apparatus of claim 13, wherein each of said at least one setting having one or more characteristics is defined with a range of values.

15. The apparatus of claim 13, wherein the database management system further stores information relating to a neural network structure associated each of the at least two sets of neural network coefficients.

16. The apparatus of claim 15, wherein the neural network structure is one of a convolutional neural network, a feed forward neural network, a neural Turing machine, Hopfield neural network, or a Boltzmann machine neural network.

17. The apparatus of claim 13, wherein the setting is one of environment, condition, and situation in which the machine operates.

18. The apparatus of claim 13, the database management system is configured to match the first data with one of at least one setting having one or more characteristics.

19. The apparatus of claim 18, wherein the database management system is configured to compare the first data with the at least one setting having one or more character-

41

istics defined with a range of values and to identify the selected one of the at least one set among one or more ranges of values that has the first data fall within its ranges of values.

20. The apparatus of claim 13, wherein the machine controlled is one of a robot, a vehicle, or a drone.

21. The apparatus of claim 13, wherein the information relating to the at least two sets of neural network coefficients is stored in a standardized format to allow access by electronic devices manufactured by different manufacturers.

22. The apparatus of claim 13, wherein the database management system further stores a set of one or more input range values associated each of the at least two sets of neural network coefficients and the instantiated neural network with the selected one from the at least two sets of neural network coefficients further configured to receive first data, and wherein the database management system further includes a trigger event detector arranged to compare the first data with the one or more input range values associated with the selected one from the at least two sets of neural network coefficients and to send a signal to the controlling device to select a new set among the at least two sets of neural network coefficients if the first data is outside the input range values.

23. The apparatus of claim 13, wherein the database management system further stores a set of one or more output range values associated each of the at least two sets of neural network coefficients, and further includes a trigger event detector arranged to compare the output with the one or more output range values associated with the selected one from the at least two sets of neural network coefficients and to send a signal to the controlling device to select a new set among the at least two sets of neural network coefficients if the output is outside the output range values.

24. An apparatus for controlling a machine, comprising: a database management system stored with at least two sets of neural network coefficients being different from each other, at least one setting having one or more

42

characteristics of a setting, and each of the at least two sets of neural network coefficients being associated with the at least one setting having one or more characteristics; and

means for, coupled to receive first data from one or more input devices of the machine, selecting one from the at least two sets of neural network coefficients based on the first data and at one least one setting having one or more characteristics, and instantiating a neural network with the selected one from the at least two sets of neural network coefficients, wherein the neural network is configured to generate an output being used to control an aspect of the machine.

25. The apparatus of claim 24, wherein each of at least one setting having one or more characteristics is defined with a range of values.

26. The apparatus of claim 24, the database management system further stores information relating to a neural network structure associated each of the at least two sets of neural network coefficients.

27. The apparatus of claim 26, wherein the neural network structure is one of a convolutional neural network, a feed forward neural network, a neural Turing machine, Hopfield neural network, or a Boltzmann machine neural network.

28. The apparatus of claim 24, the database management system is configured to match the first data with one of at least one setting having one or more characteristics.

29. The apparatus of claim 28, the database management system is configured to compare the first data with the at least one setting having one or more characteristics defined with a range of values and to identify the selected one of the at least one set among one or more ranges of values that has the first data fall within its range of values.

30. The apparatus of claim 24, wherein the information relating to the at least two sets of neural network coefficients is stored in a standardized format to allow access by electronic devices manufactured by different manufacturers.

* * * * *